

# THÈSE

Pour obtenir le grade de  
**Docteur**

Délivré par **UNIVERSITE DE MONTPELLIER**

Préparée au sein de l'école doctorale GAIA  
Et de l'unité de recherche AGAP

Spécialité : **BIDAP - Biologie, Interactions, Diversité  
Adaptative des Plantes**

Présentée par **Maxime MERCIÈRE**

**Diversité et bases moléculaires de  
l'agressivité de *Ganoderma boninense*,  
agent causal de la pourriture basale du  
stipe chez le palmier à huile**

Soutenue le 17 décembre 2015 devant le jury composé de

M. Thierry ROBERT, MC, Université Paris XI	Rapporteur
M. Phillipe SIMONEAU, PR., Faculté des Sciences d'Angers	Rapporteur
M <sup>me</sup> Claire NEEMA, PR., Montpellier SupAGRO	Examineur
M. Cyril DUTECH, CR., UMR BIOGECO INRA Bordeaux	Invité
M. Jean-Marc BOUVET, DR, UMR AGAP CIRAD Montpellier	Directeur de thèse
M <sup>me</sup> Létizia CAMUS-KULANDAIVELU, CR., UMR AGAP CIRAD Montpellier	Co-Encadrant
M. Frédéric BRETON, CR., UMR AGAP CIRAD	Co-Encadrant



## **Remerciements**

Je souhaite tout d'abord remercier mes encadrants Frédéric BRETON et Létizia CAMUS-KULANDAIVELU qui sont tous deux à l'initiative du projet SEQGANO sans lequel cette thèse n'aurait pu voir le jour. Ils m'ont accompagné et soutenu durant toute la thèse et surtout les moments très durs comme les sorties de boîte, comme à Penang, au cours de nos missions en Malaisie et en Indonésie.

Je remercie également mon directeur de thèse Jean-Marc BOUVET d'avoir bien voulu m'accepter comme thésard.

.Je remercie aussi les stagiaires que j'ai encadrés au cours de ma thèse : Romain BOULORD et Manon NOREST, avec qui j'ai passé d'agréables moments et m'ont permis de développer des compétences d'encadrant bien cachées. Je remercie aussi les autres stagiaires que je n'ai pas encadrés et qui ont été accueillis au sein de l'équipe, mais avec qui j'ai passé de très bon moment : Ian FAYOS, Emilie SICARD, Astride HONVAULT et bien d'autres.

J'adresse un remerciement spécial à Ian FAYOS pour m'avoir bien piraté mes séances de travail le week-end après m'avoir converti à CSGO.

Je remercie aussi TAN Joon Sheong pour l'aide qu'il m'a apporté au cours de ma thèse et m'avoir fait découvrir les Xiao long bao ainsi que les soirées de congrès à San Diego (I will always remember the gala dinner with you at PAG conference in San Diego).

Je remercie bien évidemment Roselyne LANNES (dite la maman de l'équipe) pour le soutien dont elle a fait preuve au cours de ma thèse.

Je remercie également toute l'équipe du bâtiment 1 : Najate MAGHNAOUI, Cathy CARASCO-LACOMBE, Marie DENIS, Virginie RIOU et Sébastien TISNE pour leur soutien morale et les repas autour de conversation sans jamais aucun sous-entendu.

Je remercie également tous mes amis (pas besoin de les citer ils se reconnaîtront ^^, et je commence à avoir des crampes dans les doigts ^^) de leur soutien au cours de la thèse (et surtout dans les derniers jours).

Je remercie mes parents pour tout ce qu'ils ont fait jusqu'à présent pour moi et m'ont conduit là où je suis aujourd'hui.

Et bien sûr la meilleur pour la fin, Je remercie ma futur femme Diane DESSENT (peut être futur MERCIERE si elle accepte de prendre mon nom ;-P) pour tout le soutien au quotidien qu'elle m'a procuré au cours de cette thèse (et avant aussi bien sur) et également sa présence et son soutien dans tous les moments durs (stress, humeur variable,...) qui ont accompagnés ces derniers mois.

# **SOMMAIRE**

## **CHAPITRE I : INTRODUCTION GENERALE .....4**

1. Apport de la génétique des populations et des nouvelles techniques de biologie moléculaire à l'étude des agents phytopathogènes .....	5
1.1. La phytopathologie : un domaine pluridisciplinaire .....	6
1.2. L'apport de la génétique des populations dans l'étude des agents phytopathogènes.....	6
1.2.1. Concept de génétiques des populations.....	6
1.2.2. La génétique des populations comme outils de la phytopathologie .....	7
1.3. L'apport de la biologie moléculaire et des techniques de séquençages pour l'étude des champignons phytopathogènes.....	9
2. Le pathosystème <i>Elaeis guineensis</i> / <i>Ganoderma boninense</i> .....	14
2.1. L'hôte : le palmier à huile <i>Elaeis guineensis</i> jacq. ....	14
2.1.1. Histoire et chiffres.....	14
2.1.2. Biologie et morphologie.....	15
2.1.3. Cycle de culture.....	16
2.1.4. L'huile .....	18
2.1.5. Les maladies en plantation industrielle .....	18
2.2. La pourriture basale du stipe causée par <i>Ganoderma boninense</i> .....	22
2.2.1. Origine et impact de la BSR .....	22
2.2.2. Biologie et diversité du pathogène.....	24
2.2.2.1. Une identification problématique.....	24
2.2.2.2. Dispersion et mode d'infection .....	25
2.2.3. Méthodes de contrôle de la maladie au champ .....	26
2.2.3.1. Pratiques culturales .....	26
2.2.3.2. Lutte chimique vs lutte biologique.....	27
2.2.3.3. Sélection variétale.....	28
3. Les objectifs de la thèse.....	29

## **CHAPITRE II : MATERIEL ET METHODES .....31**

1. Echantillonnage des isolats de <i>G. boninense</i> .....	32
2. Séquençage du génome .....	34
2.1. Optimisation du protocole d'extraction d'ADN .....	34
2.2. Choix du protocole de séquençage.....	35
3. Développement des marqueurs moléculaires .....	37

## **CHAPITRE III : ETUDE DE LA DIVERSITE GENETIQUE DE *GANODERMA BONINENSE* .....52**

1. Diversité et histoire démographique des populations de <i>Ganoderma boninense</i> au sein des plantations de Sumatra et de Malaisie Péninsulaire .....	54
2. Structure mondiale des populations de <i>Ganoderma</i> en plantation de palmier à huile : résultats préliminaires.....	81

## **CHAPITRE IV : BASE TRANSCRIPTOMIQUE DE L'AGRESSIVITE DE *GANODERMA BONINENSE* (étude préliminaire) .....83**

1. Introduction.....	84
2. Expérimentation en préépinière .....	85
2.1. Matériel et Méthodes .....	85
2.1.1. Choix des isolats candidats.....	85
2.1.2. Plan d'expérience.....	85
2.1.3. Extraction d'ARN.....	88
2.2. Résultats.....	88
2.3. Discussion .....	88
3. Expérimentation en milieu nutritif stérile .....	89
3.1. Matériel et Méthodes .....	89
3.1.1. Plan d'expérience et choix des isolats .....	89
3.1.2. Extraction d'ARN.....	90
3.1.3. Séquençage des transcrits .....	90
3.1.4. Assemblage et analyse des données RNAseq.....	90
3.2. Résultats et Discussion .....	90
3.2.1. Obtention d'ARN pur .....	90
3.2.2. Premières analyses des résultats de séquençage .....	91
3.3. Conclusion .....	93

## **CHAPITRE V : DISCUSSION GENERALE ET PERSPECTIVES.....94**

1. Développement de marqueurs microsatellites et caractérisation de trois populations distinct en Asie du Sud-Est .....	95
1.1. Une dispersion longue distance de <i>G. boninense</i> .....	96
1.2. Une expansion démographique ancienne.....	97
2. Etude préliminaire de gènes différentiellement exprimés .....	98

## **ANNEXES..... 99**

## **BIBLIOGRAPHIE .....109**



# **CHAPITRE I :**

## **INTRODUCTION GENERALE**

## **1. Apport de la génétique des populations et des nouvelles techniques de biologie moléculaire à l'étude des agents phytopathogènes**

L'observation des maladies chez les plantes remonte aux origines de l'agriculture. De par la création de l'agriculture, l'homme créa les conditions favorables à l'émergence de maladies sur les plantes cultivées, en concentrant des populations de végétaux dans des espaces restreints et en accumulant les cycles successifs de culture de la même plante sur un même sol. Depuis l'invention de l'agriculture, l'homme a cherché à comprendre l'origine de ces maladies afin de prévenir leurs apparitions ou de les combattre. Elles ont d'abord été attribuées à des causes divines, jusqu'à l'établissement d'une relation entre les facteurs climatiques, la marche des corps célestes et les maladies des plantes par Théophraste (Grèce antique, -322 à -288) (Lepoivre 2003). Ces théories ont perduré jusqu'au XVIII<sup>e</sup> siècle et l'association de la carie des céréales à un « microbe » par Isaac-Bénédict Prévost en 1807. Ce n'est qu'en 1861, avec la publication des travaux d'Anton De Bary démontrant l'association entre *phytophthora infestans* et le mildiou de la pomme de terre (Drenth 1994), que la phytopathologie est déclarée discipline scientifique. L'apparition du concept moderne d'agent pathogène se développera avec la publication des travaux de Koch en 1876 et l'établissement des fameux postulats de Koch permettant de prouver qu'un organisme isolé dans une plante infectée est bien la cause de la maladie.

À travers les âges, les maladies des plantes ont eu des effets dévastateurs sur les populations humaines. La plus représentative restera l'épidémie de mildiou de la pomme de terre, due à l'agent pathogène *Phytophthora infestans*, entre 1845 et 1852 responsable de la grande famine Irlandaise qui fera plus d'un million de morts et provoquera l'immigration de plus de deux millions d'Irlandais en majorité vers les USA (Fry 2008).

De nos jours, la lutte contre les parasites des plantes reste un enjeu majeur pour la recherche en agriculture. Les pertes économiques liées à ces parasites sont importantes et une étude les estime entre 10 et 30% du potentiel mondial de production agricole selon les cultures et les continents considérés (Agrios 2005). Dans le contexte actuel de l'agriculture, c'est-à-dire le respect de l'environnement et la réduction de l'impact de l'activité humaine sur la planète et sa biodiversité, il est important de développer des méthodes de lutte durable. Considérant ces objectifs, l'utilisation de molécules chimiques (fongicides, pesticides, ...) et la mise en œuvre de pratiques culturales à fort impact environnemental doivent être réduites. Ainsi de meilleures connaissances sur l'écosystème dans lequel se situe la culture, de la biologie de la plante et de ses pathogènes (interaction hôte/pathogène, diversité et dispersion de l'agent pathogène, ...) doivent permettre le développement de méthodes plus raisonnées et respectueuses de l'environnement.

Cette thèse s'intègre dans un projet de recherche s'articulant autour de la maladie la plus dévastatrice du palmier à huile : la pourriture basale du stipe. Ce projet vise à mieux comprendre les interactions au sein du pathosystème palmier à huile/*Ganoderma boninense*. Il intègre également la réalisation d'un « état des lieux » de la diversité du pathogène au sein des zones de culture du palmier à huile avec un intérêt particulier pour

l'Asie du Sud-Est, foyer principal de la maladie. Ce projet fait l'objet d'études regroupant quatre domaines majeurs : la génétique des populations, la biochimie, la génomique et la transcriptomique. Dans le cadre de cette thèse, seules les approches de génomique, de génétique des populations et de transcriptomique seront abordées dans le but de répondre aux problématiques du projet de recherche.

### **1.1. La phytopathologie : un domaine pluridisciplinaire**

La **Phytopathologie** ou **Phytiatrie** est la science qui traite des maladies des plantes (Lepoivre 2003). Cette discipline fait appel à des notions recouvrant plusieurs domaines : la botanique, la microbiologie, la biologie moléculaire, la génétique, la biochimie, la physiologie végétale, l'écologie, la phytotechnie, la toxicologie, l'épidémiologie et l'économie.

La phytopathologie est étroitement liée à l'exploitation d'espèces végétales (agriculture, horticulture ou sylviculture), et afin de comprendre son intérêt, il est essentiel de comprendre ce qu'est une maladie. Si l'on considère que l'exploitation d'une espèce végétale correspond à l'introduction de génotypes particuliers de végétaux (variétés ou cultivars) dans un environnement écologique particulier, le concept de maladie se rapporte alors aux anomalies observées par rapport au phénotype attendu (Lepoivre 2003). Ces anomalies portent le nom de **symptôme**. Les maladies peuvent être occasionnées par des agents phytopathogènes de nature diverse : virus, bactéries, champignons, nématodes.

Au cours d'une étude de phytopathologie il est important d'étudier les processus inducteurs de la maladie qui aboutissent à l'expression des symptômes (pathogenèse). La science qui étudie les causes des maladies porte le nom d'étiologie et constitue l'approche de base en phytopathologie, permettant l'établissement de diagnostic pertinent en vue du développement de moyen de lutte adéquate contre une maladie. La lutte contre les maladies regroupe quatre grands domaines d'action que sont : les pratiques culturales, l'amélioration génétique, la lutte chimique et la lutte biologique.

Ces quatre domaines dépendent des données biologiques recueillies à la fois sur l'hôte et mais surtout sur l'agent pathogène. Ainsi des informations sur sa physiologie, son mode de dispersion, les modalités de son interaction avec l'hôte, ainsi que sa diversité permettent le développement de nouvelles techniques dans ces quatre domaines.

### **1.2. L'apport de la génétique des populations dans l'étude des agents phytopathogènes**

#### **1.2.1. Concepts de génétiques des populations**

La génétique des populations est une discipline née au début du XXème siècle qui a permis la synthèse de deux grandes théories : la théorie darwinienne de l'évolution et la théorie mendélienne de l'hérédité. Elle a pour but l'étude de la distribution et de l'évolution au cours du temps des fréquences alléliques et génotypiques au sein d'une population ou d'un ensemble de populations interdépendantes, en fonction de différentes forces

évolutives (dérive génétique, mutation, sélection et migration) et sous l'effet des régimes de reproduction (Thomas et al. 2010).

La population peut être définie comme un ensemble d'individus dont on pense qu'il fonctionne comme une unité du point de vue évolutif et écologique (Roughgarden 1979).

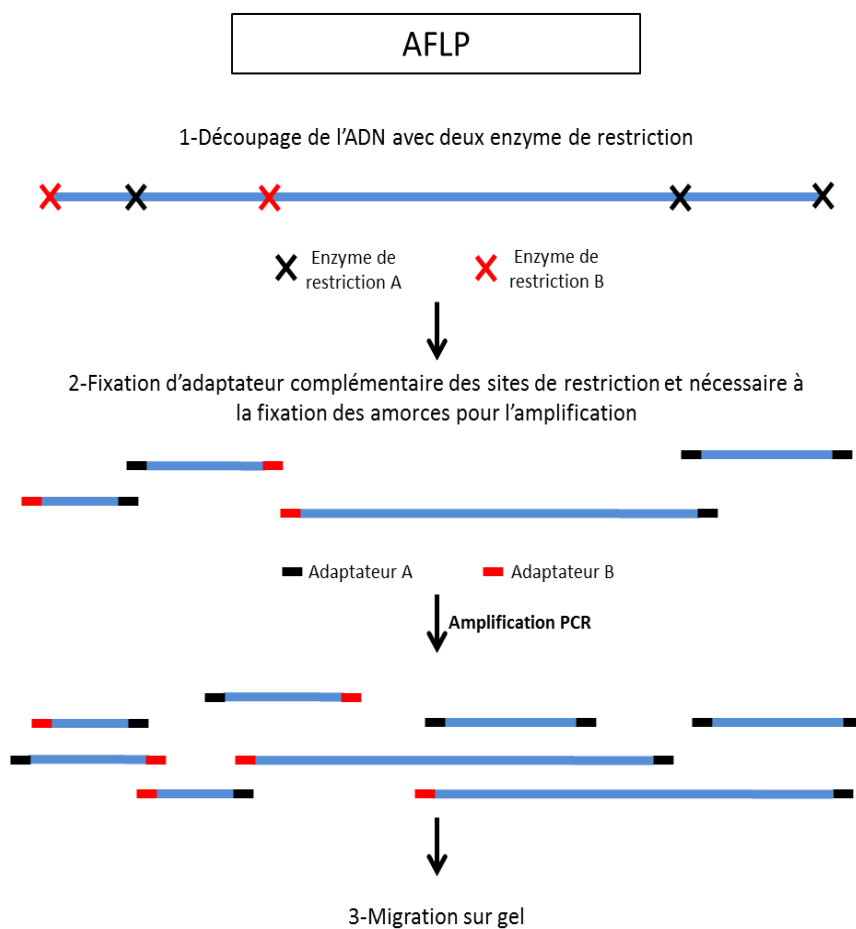
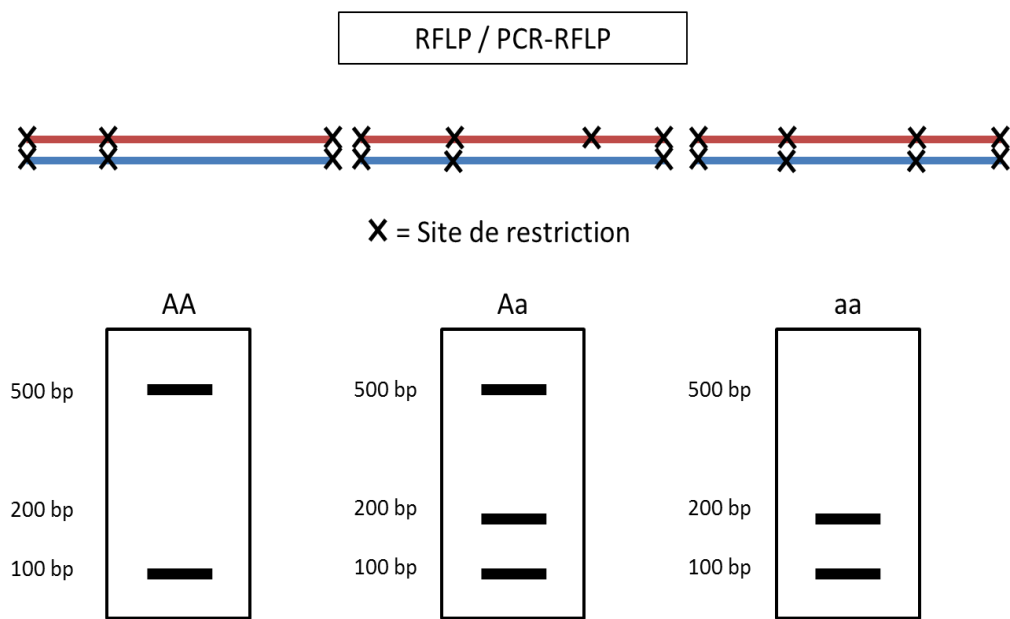
La génétique des populations s'est d'abord développée sur la base de modèles théoriques, qui ont par la suite été améliorés et validés par des approches empiriques (Hedrick 2011). Le premier modèle développé en génétique des populations et constituant le postulat de base de nombreux modèles développés aujourd'hui, est le modèle de Hardy-Weinberg (1908). Ce modèle fut établi de façon indépendante par deux scientifiques : un mathématicien britannique du nom d'Hardy, et un physicien allemand du nom de Weinberg. Ce modèle postule que : au sein d'une population, il y a équilibre des fréquences alléliques et génotypiques d'une génération à l'autre. Ce modèle est bien sûr soumis à certaines hypothèses : population de taille infinie, absence de mutation, absence de migration, pas de sélection, reproduction sexuée (espèce diploïde), panmixie (croisement aléatoire des individus : pangamie, ségrégation aléatoire des gamètes au cours de la méiose, rencontre aléatoire des gamètes ou formation aléatoire des couples) et non chevauchement des générations.

Ces hypothèses sont rarement respectées en conditions naturelles et l'observation de l'écart à l'équilibre de ce modèle sous-entend l'influence d'une ou plusieurs forces évolutives sur la structuration et l'évolution génétique de la population étudiée.

#### 1.2.2. La génétique des populations comme outils de la phytopathologie

La mise en place d'un écosystème agricole contraint l'environnement à des changements incluant : l'introduction de plante résistante aux maladies, l'utilisation d'intrants ou de fongicides, l'installation de système d'irrigation, ainsi que la rotation des cultures. Ces changements imposent une sélection forte aux populations de micro-organismes phytopathogènes, les contraignant à constamment s'adapter aux changements de leur environnement pour survivre. Cette adaptation se traduit par une évolution génétique des individus, qui transmettent leurs caractéristiques à l'ensemble d'une population par le biais de croisements au cours du cycle de reproduction sexuée ou par transfert horizontal de gènes (Friesen et al. 2006). Ainsi une lutte efficace contre les maladies des plantes cultivées passe par des stratégies de contrôle des populations d'agents pathogènes en amont des traitements (McDonald 1997).

L'étude de la composition génétique des populations et des facteurs influençant leur variation relève du domaine de la génétique des populations. L'étude de la génétique des populations d'un agent pathogène permet la compréhension de la structure génétique de ces populations, permettant le développement de stratégies de contrôle de la maladie plus efficace. La structure génétique d'une population repose sur une collection de fréquences de différents génotypes ou de fréquences d'allèles à différents loci. La répartition des fréquences génotypiques et alléliques au sein d'une population est le résultat de processus



**Figure 1. Principes de fonctionnement des marqueurs RFLP et AFLP**

évolutifs individuels couplés à des événements démographiques intra ou inter populationnels. Les processus évolutifs individuels, générant une modification des génotypes, peuvent impliquer des phénomènes de mutation, de recombinaison, de sélection et de régime de reproduction (allogamie/autogamie). De même les événements démographiques faisant varier le nombre d'individus de chaque génotype au sein de la population sont l'œuvre de phénomènes d'expansion (augmentation de la taille de la population), de migration (transfert d'un individu dans une population voisine), d'isolement par la distance (évolution indépendante d'une population par rapport aux populations environnantes) ou de goulot d'étranglement (réduction de la taille de la population par des phénomènes environnementaux). Enfin, la variabilité des populations peut être influencée par des processus aléatoires comme la fluctuation aléatoire des fréquences génotypiques pouvant créer un phénomène de dérive génétique.

Ainsi, l'étude de ces phénomènes au travers de la génétique des populations va nous permettre de mettre en évidence les facteurs influençant le plus l'évolution du pathogène et la composition génétique de ses populations, afin de déterminer le potentiel évolutif des populations de pathogènes (McDonald 1997).

### **1.3. L'apport de la biologie moléculaire et des techniques de séquençages pour l'étude des champignons phytopathogènes**

Les débuts de l'expérimentation en génétique des populations sont essentiellement basés sur l'observation de marqueurs phénotypiques, souvent peu nombreux et peu souvent mendéliens. La découverte par Oswald Avery en 1944 de l'ADN comme support du code génétique chez la bactérie du pneumocoque et la proposition d'un modèle représentant l'ADN sous forme de double hélice par James Watson et Francis Crick en 1953, ont amorcé l'essor de la biologie moléculaire. Les études de biologie moléculaire ont permis la découverte de marqueurs de la diversité moléculaire des individus d'une même espèce. Ces marqueurs d'abord basés sur des variants protéiques de type isozymes (ou allozymes) sont faciles d'étude mais présentent un polymorphisme restreint et sont peu représentatifs de la variabilité à l'échelle du génome. La découverte des enzymes de restriction par Hamilton Smith et Daniel Nathans en 1970 a permis le développement d'un nouveau type de marqueur moléculaire basé sur l'étude de la molécule d'ADN, les marqueurs RFLP (Restriction Fragment Length Polymorphism). Ces marqueurs sont co-dominants, locus spécifique et nécessitent une quantité importante d'ADN. Ils génèrent un ensemble de fragments d'ADN de longueurs différentes (dont certains fragments de taille identique peuvent être générés par des sites de restriction différents), et sont très difficiles à analyser de par le nombre de fragments d'ADN qu'ils génèrent. De plus la variation de ces fragments au niveau populationnel est souvent très faible (deux états uniques par site de restriction). L'invention de la réaction en chaîne de polymérase ou « polymerase chain reaction » (PCR) par Kary Mullis en 1983 a permis le développement de nombreux marqueurs moléculaires basés sur l'amplification de fragments d'ADN : AFLP, RAPD, minisatellite et microsatellite (Figure 1 et 2). De nos jours, les marqueurs les plus couramment utilisés en biologie

RAPD

## 1-Fixation aléatoire des amorces

## ADN de l'individu

A horizontal black bar representing a beam. Below the bar, there are four red arrows pointing to the right, each labeled 'amorce' in black text.

## 2-Amplification des brins d'ADN

### 3-Migration des amplicons sur gel

**amplicons**

SSR
-----

Individu 1  
allèle A1

Primeur F1

ATATATATATATATATATATATAT

Primeur R1

Individu 2  
allèle A2

Primeur F1

ATATATATATATATAT

Primeur R1

### Amplification PCR

## Lecture allèles sur gel

Individu 1

Individu 2

Différence de 12  
nucléotides

**Figure 2. Principes de fonctionnement des marqueurs RAPD et SSR**

moléculaire et en génétique des populations sont les marqueurs microsatellites (Jarne et al. 1996, Schoebel et al. 2013 a, Presti et al. 2014). Ces séquences d'ADN répétées, dont les motifs sont formés de 1 à 6 nucléotides, sont distribuées le long du génome, plus particulièrement dans les régions non codantes (marqueur neutre, n'influençant ni le caractère étudié, ni la valeur sélective). Ce sont des marqueurs co-dominants, hautement polymorphes et particulièrement abondants dans la plupart des génomes eucaryotes et procaryotes. Ces marqueurs sont obtenus à partir de la construction de banques d'ADN enrichies en séquence microsatellite et l'emploi de technologie de séquençage Sanger, ce qui est une opération longue et coûteuse (Zane et al. 2002). Le principal défaut de ces marqueurs réside dans la possible génération d'allèles nuls due à une mutation ponctuelle ou une insertion/délétion (Indel) dans les régions de fixation des amorces. Ces allèles nuls sont difficiles à mettre en évidence, mais certaines méthodes utilisées en génétique des populations, permettent de les détecter (Dabrowski et al. 2014) et peuvent éviter le séquençage de la séquence répétée et des séquences bordantes correspondant aux sites de fixation des amorces pour vérifier la présence d'une mutation ponctuelle (méthode la plus fiable).

Les développements de la première technique de séquençage, présentée par Frederick Sanger (1977), couplée à la technique de PCR, ont conduit à la course au séquençage du plus grand nombre d'espèces possibles (<http://www.1000genomes.org>). Le principe de séquençage développé par Sanger est basé sur l'utilisation de didésoxyribonucléotides et d'un gel d'électrophorèse. L'utilisation de didésoxyribonucléotides permet l'interruption aléatoire de la synthèse enzymatique des brins d'ADN complémentaires et est appelée technique de terminaison de chaîne. L'évolution des techniques de séquençage a permis la création de nouvelles disciplines (génomique, transcriptomique) et le développement d'un autre type de marqueur : les SNP (Single Nucleotide Polymorphism). Ces marqueurs sont répartis de façon relativement uniforme le long des génomes et sont hautement informatifs sur l'évolution des génomes malgré leur faible polymorphisme (en général, seuls deux états sont possibles). Afin de compenser ce faible niveau de polymorphisme, il est nécessaire d'en utiliser un très grand nombre en comparaison des marqueurs microsatellites. Jusqu'à présent les études de diversité basées sur l'utilisation de marqueurs SNP n'étaient pas majoritaires du fait de leur prix élevé lié à une étape de développement nécessitant un nombre de séquences important afin de les valider. Ces marqueurs sont utilisés principalement sur des espèces dont la séquence du génome est disponible et de bonne qualité. Les marqueurs microsatellites restent donc plus abordables notamment grâce à l'évolution des techniques de séquençage qui ont permis l'abaissement des coûts de séquençage (Schoebel et al. 2013 b).

En 1988, Ronaghi et Uhlen développent une nouvelle méthode de séquençage basée sur l'ajout en temps réel d'un seul nucléotide, lui-même révélé en temps réel par détection de la luminescence libérée par l'incorporation du nucléotide (Hyman 1988). Cette technique est appelée pyroséquençage. Le XXI<sup>e</sup> siècle verra l'avènement de la nouvelle génération de



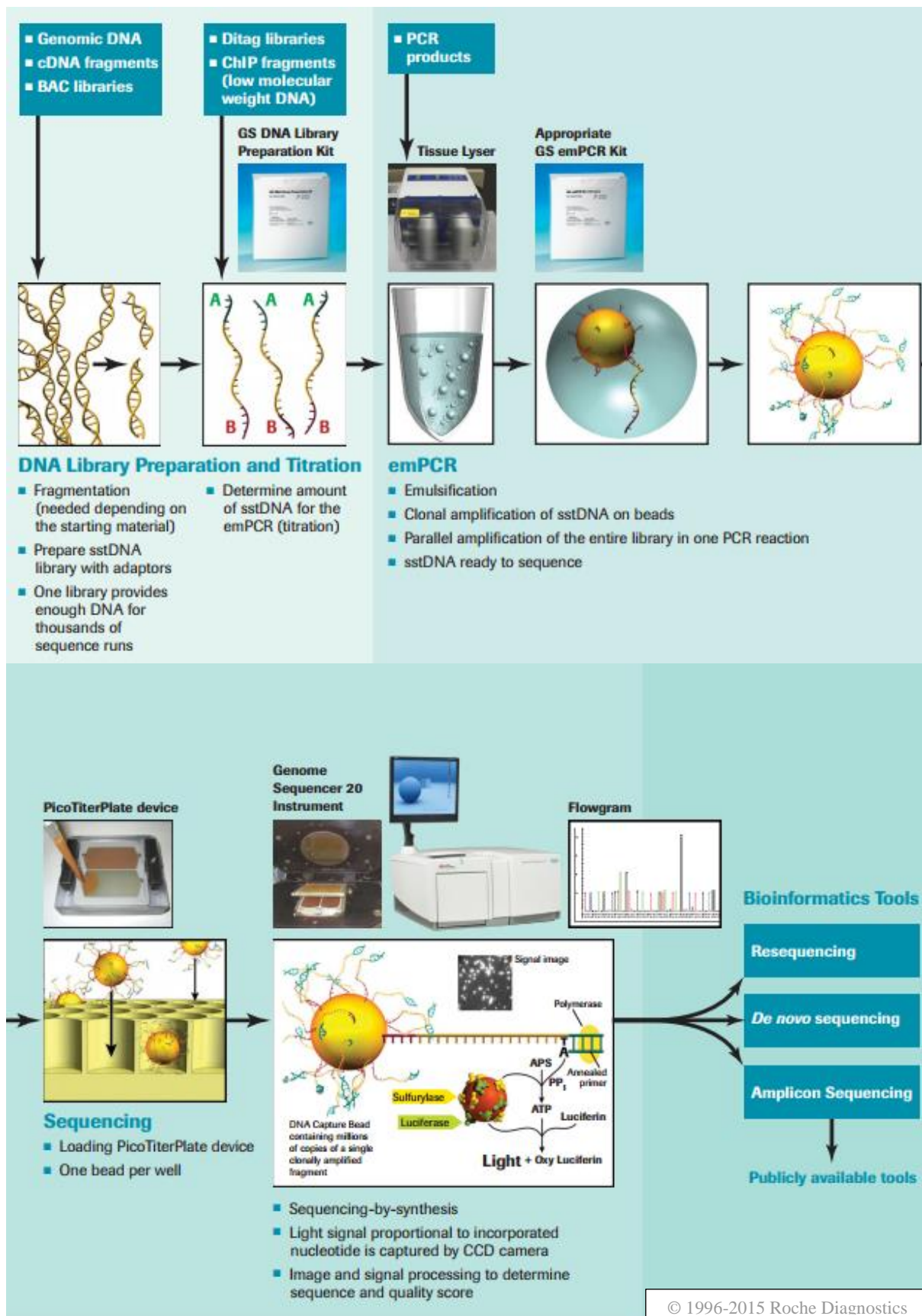
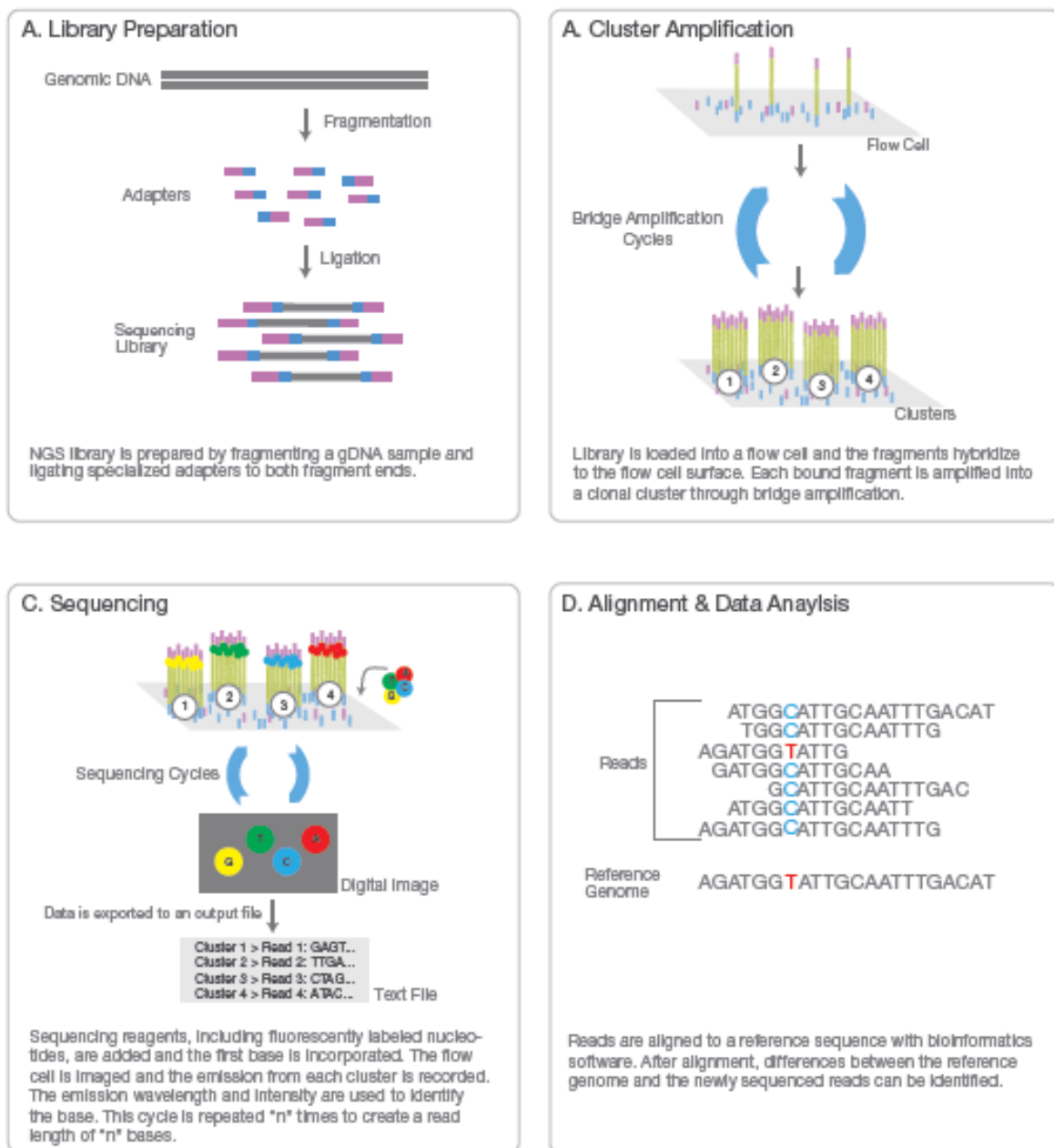


Figure 3. Principe du séquençage 454 Roche.



**Figure 4. Principes du séquençage illumina**

techniques de séquençage, qui seront des adaptations des techniques de séquençage Sanger et de pyroséquençage. Ces nouvelles méthodologies de séquençage permettant un séquençage haut débit, deviennent possibles grâce aux avancées technologiques issues de l'évolution des connaissances en physique, informatique, chimie, nanotechnologie et biotechnologie. Deux technologies majeures sont aujourd'hui utilisées : la technologie 454 commercialisée par la société 454 life science et la technologie CRT Solexa/Illumina.

La technologie 454 est basée sur le principe du pyroséquençage couplé à un ensemble de technologies : les plaques constituées de fibre optique picotitrée qui contiennent 1,6 million de puits, la PCR en émulsion (emPCR) dans des microréacteurs (gouttelettes d'une émulsion) pouvant réaliser 300 000 réactions en parallèle, ainsi que des

technologies informatiques de pointe pour l'acquisition, le traitement et l'analyse des images. Les différentes étapes du processus de séquençage sont décrites dans la figure 3. Cette technologie employant la nanotechnologie permet de traiter plus de 20 millions de bases nucléotidiques par cycle de quatre heures. Cependant, les séquences obtenues par la technologie 454 sont en moyenne de 500 paires de bases (pb) et sont relativement courtes par rapport aux séquences obtenues par séquençage Sanger classique.

La technologie CRT Solexa/Illumina est la seconde des deux techniques actuellement utilisées. Elle est basée sur une méthodologie dérivée du séquençage Sanger appelé CRT (cyclic reversible termination), qui est basée sur le même principe de terminaison de chaîne mais qui est réversible dans cette technologie. Cette méthode a été adaptée pour une technologie de séquençage sur puce et intègre plusieurs technologies : les biopuces à ADN, la nanotechnologie, ainsi que des technologies informatiques de pointe pour l'acquisition, le traitement et l'analyse des images. Les étapes constituant cette méthode sont décrites dans la figure 4. Cette technologie génère des séquences plus courtes (environ 150 bp) que la technologie 454 mais cette courte longueur de séquence est compensée par un nombre de séquences 100 fois supérieur à la technologie 454 (environ 3 milliards de séquences générées par cycle).

L'amélioration des techniques de séquençage a non seulement permis la découverte de nouveaux marqueurs, l'amélioration du développement des marqueurs microsatellites, mais a rendu également accessible l'étude intégrale des génomes, ainsi que l'étude simultanée des transcriptomes de plusieurs individus. Ces approches ouvrent de nouvelles perspectives dans l'étude des voies de signalisation régissant l'interaction hôte/pathogène.

## **2. Le pathosystème *Elaeis guineensis*/*Ganoderma boninense***

### **2.1. L'hôte : le palmier à huile *Elaeis guineensis* Jacq.**

Le palmier à huile (genre *Elaeis*) est une monocotylédone tropicale, arborescente, de la famille des Arécacées. Les palmiers à huile d'intérêt économique regroupent deux espèces *Elaeis guineensis* Jacq. et *Elaeis oleifera*. *Elaeis guineensis* Jacq. est une espèce provenant d'Afrique, principalement exploitée en plantation commerciale sous forme de monoculture. Cette espèce est à l'origine de l'expansion actuelle de la culture du palmier à huile en Asie du Sud-Est et en Afrique. La seconde espèce, *Elaeis oleifera*, n'est pas cultivée commercialement en tant que telles, mais son hybride avec *Elaeis guineensis* est exploité en Amérique latine en raison de sa tolérance à une maladie mortelle majeure présente en Amérique latine : la pourriture du cœur (Jacquemard 2011). Au travers de ce mémoire nous nous focaliserons sur *Elaeis guineensis* Jacq., qui est l'espèce la plus largement plantée et sur l'attaque du pathogène *Ganoderma boninense* entraînant des pertes économiques sévères.

#### **2.1.1. Histoire et chiffres**

*Elaeis guineensis* Jacq. est originaire d'Afrique tropicale et plus précisément du golfe de Guinée où subsistent encore quelques palmeraies naturelles. Cette espèce fut introduite en

Asie du Sud-Est par le biais de 4 individus issus de graines cultivées au jardin de Bogor (Indonésie) en 1848. Le début de l'exploitation industrielle d'*E. guineensis* à des fins d'exportation remonte au début du XX<sup>ème</sup> siècle (1911 pour l'Indonésie et 1917 pour la Malaisie) (Jacquemard 2011). Elle est devenue un moteur économique et de développement important dans de nombreux pays d'Asie du Sud-Est, qui combinent fortes densités de population et taux de croissance économiques élevés. Le palmier à huile est aujourd'hui la première plante oléagineuse mondiale avec un totale de 54 millions de tonnes d'huile produites en 2012 (département d'agriculture des Etats-Unis d'Amérique: <http://apps.fas.usda.gov/>) dont 80% est utilisé en industrie alimentaire. En regard du marché mondial des huiles végétales, cette production représente 32% de la production annuelle pour seulement 5,5% de la surface de terre utilisées pour la production d'huile végétale. Actuellement, la culture du palmier à huile se répartit strictement dans les zones intertropicales: l'Asie du Sud-Est, l'Afrique et L'Amérique Latine. Cependant en 2013, deux pays, l'Indonésie (50 % de la production mondiale et 28,4 millions de tonnes) et la Malaisie (34% de la production mondiale et 19,2 millions de tonnes), concentraient à eux seuls 84 % de la production mondiale d'huile de palme. Les 18% restant se partageaient entre la Thaïlande (3% et 1,9 million de tonnes), la Colombie (2% et 1 million de tonnes), le Nigeria (2% et 0,9 million de tonnes), la Papouasie/Nouvelle-Guinée (1% et 0,5 million de tonnes), l'Équateur (1% et 0,5 million de tonnes) et quelques autres pays minoritaires d'Afrique et d'Amérique Latine représentant 7% de la production mondiale (données issues de la base de données Oil World : <http://www.oilworld.biz/>). Le principal atout de cette monoculture oléagineuse est sa productivité inégalée par les autres cultures de plantes oléagineuses (Colza, Tournesol et Soja), avec un rendement moyen de 4 tonnes à l'hectare et par an, contre moins d'une tonne à l'hectare et par an pour les autres principales plantes oléagineuses. Les principaux consommateurs sont : la Chine avec 25,3 kg/habitant/an (1,35 milliard d'habitants), suivie de l'Union Européenne avec 60 kg/habitant/an (503,6 millions d'habitants), l'Inde avec 15 kg/habitant/an (1,26 milliard d'habitants) et les Etats-Unis d'Amérique avec 55,3 kg/habitant/an (315,8 millions d'habitants).

#### 2.1.2. Biologie et morphologie

*E. guineensis* est une plante pouvant atteindre 25 à 30 m de haut en milieu naturel mais dont la hauteur en palmeraie ne dépasse que rarement les 15 m en raison du temps de rotation entre deux générations de palmiers (de 17 à 35 ans, en moyenne 25 ans (Corley et al. 2003)) ajouté à la difficulté à récolter les régimes de fruits à de grandes hauteurs. Cette monocotylédone possède un stipe, tige ligneuse terminée par un faisceau de feuilles (<http://dico-sciences-animales.cirad.fr>) appelé couronne foliaire comportant de 30 à 45 feuilles vertes et contenant en son cœur un unique bourgeon végétatif. Le système racinaire est de type fasciculé et formé de plusieurs milliers de racines (Corley et al. 2003).

Le **stipe** se différencie du tronc des plantes dicotylédones par l'absence de croissance en épaisseur (absence de la couche de cellules méristématiques appelée cambium, située entre

le bois et le liber). Il se présente sous forme d'une colonne tronconique à la base sur une hauteur de 1 m à 1,5 m, puis présente un diamètre constant. Le diamètre à la base est de 80 à 110 cm et de 40 à 50 cm pour la partie cylindrique. Les chicots pétiolaires laissés après l'élagage des feuilles restent adhérents jusqu'à 20 ans et finissent par tomber, laissant une trace triangulaire sur le stipe.

La **couronne foliaire** comporte des feuilles irrégulièrement pennées dont la longueur peut atteindre 10m et possèdent un poids frais pouvant atteindre 8 kg. La feuille est composée d'un pétiole d'environ 1,5 m, d'un rachis pouvant atteindre plus de 7 m et de 250 à 350 folioles réparties de façon relativement égale de chaque côté du rachis.

Le **système racinaire** s'insère sur un plateau racinaire situé à la base du stipe. Ce plateau racinaire atteint environ 80 cm de diamètre et s'enfonce de 40 à 50 cm dans le sol. Les racines du palmier à huile peuvent descendre à plus de 6 m de profondeur et s'étendre sur plus de 20 m de rayon.

Le palmier à huile est une plante monoïque pour laquelle se succèdent des cycles de spadices mâles et femelles. Sa pollinisation est essentiellement entomophile et assurée par les insectes du genre *Elaeidobius*, notamment le charançon Curculionidés (*Elaeidobius kamerunicus*), originaire d'Afrique et introduit en 1984 dans toutes zones de culture à travers le monde. Enfin, l'entrée en production du palmier s'établit entre 2,5 et 4 ans et atteint un rythme de production de croisière à partir de 5 à 7 ans en fonction du fonds génétique de l'individu et des conditions de culture.

### 2.1.3. Cycle de culture :

Avant la mise en terre définitive d'un jeune plant, la production du jeune plant passe par quatre étapes décrites par Jacquemard (2011). Cette production du jeune plant peut durer entre 12 et 16 mois suivant le fonds génétique de la plante et les pratiques culturales.

- **Le Germoir**

La première étape de production du jeune plant est la germination. Elle est réalisée dans un bâtiment appelé Germoir constitué de pièces dont les conditions (température, humidité...) sont adaptées à la levée de la dormance, ainsi qu'aux autres étapes de la germination. Le processus de germination commence par la levée de la dormance en rétablissant un taux d'humidification de la graine suffisamment élevé ainsi qu'un traitement thermique pouvant durer de 60 à 80 jours. En effet, le stockage des graines implique un processus de déshydratation permettant une baisse de la teneur en eau de la graine, ainsi qu'une meilleure conservation des semences. La graine rentre ensuite dans l'étape de germination qui est constitué d'une première étape de trempage d'environ 5 jours permettant l'humidification nécessaire à la germination de la graine. Elle est ensuite transférée dans une pièce sombre à température contrôlée puis stockée en attendant la livraison des graines germées.

- **La Prépépinière**

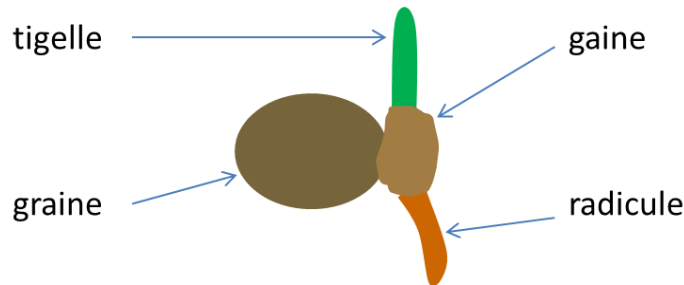


Figure 6. Schéma de graine germée utilisé en prépépinière

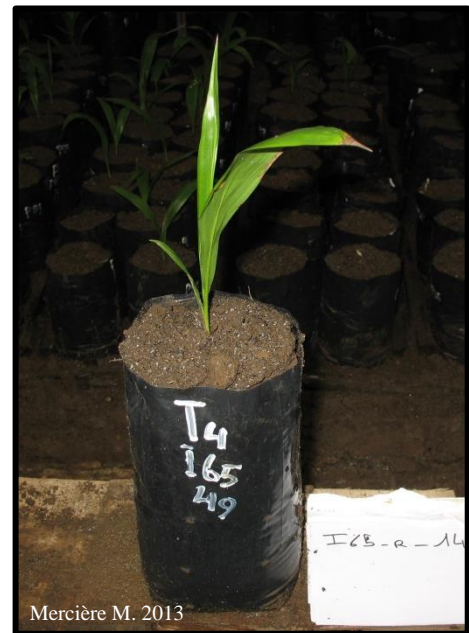


Figure 5. Photo de graine germée au stade trois feuilles en prépépinière. Les marques de flétrissement en bout des feuilles témoignent de l'infection de la plantule par *G. boninense*

La seconde étape correspond à la culture des jeunes palmiers sous ombrage pendant les 2 à 4 mois qui suivent la germination. Pendant cette étape la graine germée va être repiquée, dans un sac de plastique rempli de terre stérile et tamisée, à partir du moment où la tigelle et la radicule se sont bien différenciés (Figure 5). Au cours du premier mois, le jeune palmier va émettre ses deux premières feuilles ainsi que des racines adventives. Au bout de quatre mois le jeune palmier aura développé un système racinaire comportant des racines primaires, secondaires et tertiaires et sera devenu complètement autotrophe (capable de générer sa propre matière organique à partir d'éléments minéraux). Au cours de ces quatre mois, le jeune palmier aura également développé trois à quatre feuilles à limbe lancéolé et sera prêt pour son transfert en pépinière.

- **La pépinière**

C'est la dernière étape avant la mise en terre définitive du jeune palmier. Cette étape dure de 7 à 9 mois, pendant lesquels le palmier va perdre peu à peu son aspect juvénile et produire des palmes de taille réduite mais morphologiquement semblable aux adultes. Au cours de cette étape le palmier est toujours en terre contenue dans un sac de plastique mais l'ombrage est progressivement levé. Après cette étape de 9 mois les plants sont transférés en plantation et plantés à une densité de 143 palmiers par hectare (dispositif en triangle équilatéral de 9 m de côté, soit un espacement de 7,80 m entre lignes).

#### 2.1.4. L'huile :

À partir de la fécondation des fleurs de l'inflorescence femelle, les épis vont évoluer en régimes pouvant posséder de 500 à 3000 fruits. La période de maturation des fruits va durer de 5 mois et demi à 6 mois. À partir de ces fruits vont être extraites deux huiles différentes. La première huile, l'huile de palme, est utilisée principalement pour l'alimentation humaine (80%). Elle est extraite à partir de la pulpe (mésocarpe) du fruit par un processus impliquant un malaxage à chaud. La couleur de cette huile varie du jaune orangé clair au rouge orangé foncé. La seconde huile, l'huile de palmiste, est extraite à partir du pressage de l'amande (appelée également palmiste). Elle est principalement utilisée dans la fabrication de produits cosmétiques et de savon, se substituant à l'huile de coprah. L'huile de palme représente un total de 23 à 26 kg pour 100 kg de régime, contre en moyenne 4 kg d'huile de palmiste (Jacquemard 2011).

#### 2.1.5. Les maladies en plantation industrielle

Les maladies affectant le palmier à huile sont nombreuses, peuvent se déclarer à différents stades de développement de la plante et affecter différents organes. Certaines maladies bénignes, souvent présentes aux stades de développement précoces, (peuvent être endiguées par de simples pratiques culturales et n'ont pas de réel impact économique sur la plantation. Pour d'autres maladies plus graves, la survie de l'arbre peut être engagée et donc impacter la production d'huile. On dénombre 14 maladies répertoriées (Jacquemard 2011), dont 5 n'ont toujours pas d'agent causal identifié ou confirmé et 5 autres ont une origine fongique (5 sur 14 maladies répertoriées). Les maladies du palmier à huile sont en général regroupées suivant le stade de développement auquel elles apparaissent (Corley et al. 2003, Jacquemard 2011).

##### 2.1.5.1. Les maladies de préépinière

Deux maladies sont majoritairement observées en préépinière. La première maladie est la **fonte des semis** que l'on retrouve dans toutes les zones de culture du palmier à huile. Cette maladie est causée par divers agents non spécifiques et provoque une absence de levée. Des stratégies prophylactiques ont été développées : éviter les excès d'ombrage, d'arrosage ou les sols trop lourds (trop riches en argile) et désinfecter les sols utilisés en préépinière.

La seconde maladie, l'**Anthraxnose**, peut être causée par de nombreux champignons, principalement des ascomycètes. Il n'existe pas de pathogène spécifique identifié pour cette maladie, mais de nombreux champignons du genre *Colletotrichum* ont pu être isolés à partir de palmiers présentant les symptômes de l'anthraxnose. Les plantes atteintes de cette maladie présentent sur les feuilles des taches brunes à bords arrondis avec un centre gris et cernées d'un halo jaunâtre. La lutte contre l'anthraxnose passe par des moyens préventifs, tels que la maîtrise de l'ombrage et de l'arrosage, ainsi que la proscription des sols trop lourds. L'anthraxnose est présente dans deux zones géographiques : l'Afrique et l'Asie du Sud-Est.

##### 2.1.5.2. Les maladies de pépinière



On dénombre quatre maladies principales sévissant en pépinière : le *Blast*, la Cercosporiose et la pourriture sèche du cœur, toutes trois observées en Afrique et le *Curvularia* observé en Asie du Sud-Est.

Le **Blast** est une maladie à bactéries de type mycoplasme, transmise par l'insecte suceur *Recilia mica*, une cicadelle que l'on retrouve principalement en république du Congo, au Libéria et en Côte d'Ivoire. Elle provoque la pourriture de la flèche et du cortex racinaire et conduit dans la plupart des cas à la mort de la plante. Elle peut également être rencontrée en pré-pépinière et au cours de la première année de plantation. Des moyens de lutte préventive ont été développés pour éviter l'apparition de cette maladie essentiellement liée à la présence de graminées. Les moyens employés sont : le désherbage des pépinières et dans un rayon de 40 m autour des sites de pépinière, ainsi qu'un ombrage pendant les périodes sensibles, ...).

La **cercosporiose** est provoquée par le champignon foliaire *Cercospora elaeidis*. Elle apparaît principalement sous forme de symptômes foliaires (mouchetures brun-orange sur feuilles âgées suivies d'un dessèchement des extrémités et des bordures du limbe), n'a pas d'impact économique important en plantation et peut être rencontrée à de rares occasions en pré-pépinière. Cependant elle provoque un effet dépressif sur la croissance des plants et fait donc l'objet d'une lutte chimique par pulvérisation d'un cocktail de fongicides sur la face inférieure des feuilles où se développe le champignon.

La **pourriture sèche du cœur** est une maladie mortelle dont l'agent pathogène n'a pas encore été identifié. Cependant cette maladie a été associée à deux principaux vecteurs : *Sogatella cubana* et *Sogatella kolophon*. Les symptômes principaux sont l'apparition de taches jaunâtres sur la flèche et de part et d'autre de la nervure centrale, l'arrêt de la croissance et l'apparition de taches huileuses dans le plateau racinaire. Les moyens de lutte sont essentiellement préventifs et identiques à celles employées pour le *blast*.

La **Curvularia**, provoquée par une espèce non identifiée de champignon ascomycète du genre *Curvularia*, est une maladie foliaire se développant principalement à la suite d'un stress important subi par la plante. Elle n'a pas un effet économique significatif en plantation et peut également être rencontrée à de rares occasions en pré-pépinière. La lutte contre cette maladie passe par la suppression de la cause du stress et un traitement chimique identique à la cercosporiose.

- **Les maladies de plantation**

Les plantations étant des environnements moins contrôlés que les pré-pépinières ou les pépinières, on rencontre un nombre plus important de maladies. On dénombre aux totales 9 maladies réparties sur les trois grandes zones géographiques de culture du palmier à huile dont 6 d'entre elles sont spécifiques à l'Amérique latine.



La **fusariose** est présente principalement en Afrique (dans toutes les zones de culture) mais aussi en Amérique Latine. C'est l'une des trois grandes maladies ayant un impact économique majeur en plantation. L'agent pathogène responsable de la fusariose est *Fusarium oxysporum* pv. *elaeidis*, un champignon ascomycète que l'on peut également retrouver en pépinière. Il se répand de proche en proche par contamination racinaire et provoque de nombreux symptômes foliaires (dessèchement des feuilles moyennes, non-ouverture des flèches, cassure au premier tiers du rachis), réduit la production et provoque l'apparition de fibre brune dans le cœur du stipe. La fusariose provoque souvent la mort de l'arbre, mais des cas de rémission ont été observés. Une lutte intégrée passant par l'éradication des arbres malades et l'extirpation des souches permet de limiter la propagation de la maladie. Il existe également programme de développement de matériel tolérant afin de replanter des palmiers dans les plantations atteintes mais toujours le plus loin possible des palmiers ayant déjà été abattus et dessouchés.

La **maladie des anneaux rouges** est une maladie présente uniquement en Amérique latine. Elle est provoquée par le nématode *Rhadinaphelenchus cocophilus* qui est transmis d'arbre en arbre par le vecteur *Rhyncophorus palmarum*. Cette maladie provoque le ralentissement de la croissance, le jaunissement des plus jeunes feuilles qui restent serrées en une masse compacte et pourrissent ainsi que l'avortement partiel ou total des inflorescences ou des fruits. La maladie des anneaux rouges conduit à la mort des palmiers et seules des pratiques culturales adaptées permettent de prévenir l'apparition de la maladie : prévention des blessures, contrôle des populations du vecteur (piégeage et destruction des gîtes larvaires) en plantation et lutte contre tous les ravageurs pouvant occasionner des blessures.

La **maladie des taches annulaires** qui sévit en Amérique latine n'a pas d'agent pathogène identifié, seule une relation entre la maladie et la présence de graminées autour des arbres infectés a pu être établie. Cette maladie touche principalement les jeunes palmiers en plantation et son incidence diminue avec l'âge des arbres (à partir de 5-6 ans). C'est une maladie mortelle qui provoque la chlorose des jeunes feuilles et des taches annulaires sur les folioles, suivie d'un jaunissement de l'ensemble de la couronne foliaire et de l'apparition de taches brunes nécrotiques sur le plateau racinaire. Les seules techniques développées pour lutter contre cette maladie sont des moyens de lutte intégrée passant par l'éradication des palmiers malades, l'élimination des graminées et l'installation de plantes couvrantes du type légumineuses. Il a cependant été constaté une relative tolérance de l'hybride *E. oleifera* x *E. guineensis* à cette maladie.

La **marchitez sopressiva** est présente uniquement en Amérique latine et est une maladie à protozoaire, induite par un protozoaire flagellé intra-phloémique du genre *Phytomonas* dont l'espèce n'est pas identifiée. Ce protozoaire est transmis par des insectes piqueurs du genre *Lincus* et *Ochlerus*. Cette maladie provoque une chlorose générale des feuilles basses gagnant rapidement les feuilles hautes, ce qui conduit à un dessèchement total de la couronne foliaire en moins d'un mois, ainsi qu'un pourrissement des régimes et des racines.

C'est une maladie mortelle foudroyante pour laquelle seuls des moyens de lutte préventive ont été développés (éradication des arbres malades, contrôles des populations de vecteur en plantation). Les hybrides *E. oleifera* x *E. guineensis* sont moins sensibles à cette maladie.

La **pourriture du cœur** (PC) est une maladie mortelle décrite principalement en Amérique latine mais qui a été signalée récemment en Côte d'Ivoire (De Franqueville, pers. Com.). Les symptômes de cette maladie sont variables et aucun agent pathogène n'a été identifié comme responsable de cette maladie considérée comme l'une des trois plus importantes maladies du palmier à huile. Certains auteurs suggèrent la présence en plantation de différents types de pourriture du cœur en relation avec les observations symptomatologiques des palmiers infectés. Ceci peut sous-entendre l'implication de plusieurs agents pathogènes à des degrés divers. Cependant, de récentes études suggèrent qu'une espèce fongique du genre *Phytophthora* serait l'agent causal principal de la pourriture du cœur (Torres et al. 2010). La PC atteint les arbres dès les plus jeunes stades et se manifeste en général par des symptômes foliaires pouvant aller de la chlorose au dessèchement des palmes et provoquant un pourrissement de la couronne foliaire. La lutte contre cette maladie passe par l'élimination des arbres malades et la destruction des souches. Des études récentes en amélioration génétique réalisées par PalmElite (filiale du Cirad) ont permis de mettre en évidence la tolérance de certains hybrides *E. oleifera* x *E. guineensis* vis-à-vis de la PC.

La **pestalotiopsis** est une maladie présente en Amérique latine et est induite un champignon ascomycète du genre *Pestalotiopsis* qui attaque le palmier via les blessures causées par les insectes défoliateurs ou piqueurs. Elle se manifeste par l'apparition de taches grises sur les folioles avec un halo jaunâtre au début. La croissance des taches finit par rendre les feuilles totalement grises et cassantes, avec des folioles pouvant se nécroser entièrement. Les méthodes de lutte passent par le contrôle des populations d'insectes défoliateurs et piqueurs grâce à des pesticides et l'élagage des feuilles contaminées afin de diminuer les sources d'inoculum.

La dernière maladie est la **pourriture basale du stipe** ou **basal stem rot** en anglais (BSR), qui est causée par un pathogène fongique, *Ganoderma boninense*. Cette maladie est la plus dévastatrice pour les planteurs et touche principalement l'Asie du Sud-Est. Elle a été observée plus récemment en Afrique et Amérique latine mais semble atteindre uniquement les arbres très âgés ou déjà affaiblis par une autre maladie. Les symptômes foliaires des palmiers infectés se caractérisent par des flèches non ouvertes et un dessèchement des palmes basses. On remarque également l'apparition de sporophores sur le stipe ainsi qu'une pourriture sèche de la base du stipe conduisant à l'apparition d'une cavité plus ou moins importante. L'accroissement de la dégradation des tissus racinaires et de la base du stipe par ce pourridié va mettre en péril, entre autres, le continuum nutritionnel du palmier (racine-stipe), qui va conduire à l'apparition des symptômes foliaires et à terme à la chute de l'arbre. Les moyens de lutte contre cette maladie sont relativement limités et regroupent



**Figure 7. Photo d'une clairière due à l'impact de la BSR dans la plantation de Tanah Gambus (Indonésie)**

principalement des pratiques culturales intervenantes à différents stades de la culture du palmier à huile. Avant replantation, les palmiers malades sont détruits et leurs souches extirpées afin de réaliser un trou de sanitation. Les stipes et les souches des palmiers abattus sont déchiquetés et épandus sur des terrains inoccupés pour diminuer la charge d'inoculum à l'aide de l'ensoleillement, augmentant la température qui devient trop importante pour le champignon. Des programmes de sélection génétique ont permis le développement de variétés de palmiers au niveau de résistance variable qui sont utilisées en replantation le plus loin possible des anciennes souches de palmier contaminées. Dans le but de sauvegarder les génotypes d'intérêt majeur (présents au sein de jardins semenciers) pour les programmes d'amélioration, les tissus infectés du stipe sont éliminés (chirurgie), puis un buttage favorisant la formation de nouvelles racines est réalisé afin de sauvegarder au mieux les palmiers infectés. Des tests sur des organismes antagonistes (*Trichoderma* sp.) utilisés avant replantation ont également été lancés.

## **2.2. La pourriture basale du stipe causée par *Ganoderma boninense***

### **2.2.1. Origine et impact de la BSR**

Le basidiomycète tellurique *Ganoderma boninense*, responsable de la BSR, est considéré comme appartenant aux pourritures blanches ou fibreuses (champignon dégradant la lignine et la cellulose). C'est un champignon hemibiotrophe provoquant un pourrissement des





**Figure 8. Fructification multiple de *G. boninense* à la base d'un stipe de palmier à huile dans la plantation de Tanah Gambus (Indonésie)**

racines et de la base du stipe, ayant pour conséquence une perturbation de la circulation des nutriments jusqu'à la couronne foliaire.

L'infection de palmier à huile par différentes espèces de *Ganoderma* a été reporté pour la première fois en Afrique en 1920 (Wakefield 1920), puis dès 1928 en Malaisie (Sharples 1928). À cette période la présence de BSR en plantation n'a pas d'incidence économique importante. Le pathogène n'est signalé que sur des palmiers âgés en fin de cycle d'exploitation, considérés comme prédisposés à la maladie de par leur sénescence (Turner et al. 1974). Dès 1956, dans les plantations succédant à des cultures de cocotiers ou à une replantation, la BSR commence à se manifester sur quelques jeunes arbres de 5 ans et de façon plus importante sur des arbres de 10-15 ans (Larter 1956, Anon 1958). En 1981, la BSR s'est fortement propagée dans la majorité des plantations, de telle sorte que Turner (1981) la décrit comme « la maladie la plus dévastatrice en palmeraie, causant des pertes significatives en Asie du Sud-Est ». On estime aujourd'hui qu'elle peut infecter jusqu'à 80% des arbres d'une palmeraie (suivant les antécédents de la plantation : précédent cultural, type de sol, nombre de cycles/génération,...)(Turner 1965 c) réduisant drastiquement la production de régime et donc la production d'huile.

Cependant, la première observation de sporophores attribués à l'espèce *Ganoderma boninense* a été répertoriée en Asie du Sud-Est dans des plantations de cocotiers. Toutefois, cette présence de sporophores sur cocotiers ne s'accompagne de l'apparition d'aucun symptôme de dépérissement.

La présence antérieure de cocotiers dans certaines zones géographiques actuellement occupées par des plantations de palmier à huile semble un facteur ayant favorisé le développement de *G. boninense* sur palmier à huile. Les observations de Turner (1965 c) en plantations commerciales de palmier à huile montrent que les taux les plus bas d'infection sont obtenus pour les zones faisant suite à une culture d'hévéa (2 à 4 % de palmiers infectés). En revanche, les plantations faisant suite à une replantation ou à une plantation de cocotiers montrent des taux d'infection respectivement de 24 à 28% et 6 à 81%. Plus récemment, Gurmit (1991) a mis en évidence une augmentation de l'incidence de la maladie au fil des années d'un même cycle de culture mais aussi un effet significatif des cycles de cultures successifs sur l'apparition précoce de la maladie en replantation. Ces résultats suggèrent une présence antérieure de l'agent causal de la BSR et sous-entendent une possible adaptation de *G. boninense* au palmier à huile.

## 2.2.2. Biologie et diversité du pathogène

### 2.2.2.1. Une identification problématique

Plusieurs espèces de *Ganoderma* ont été mises en évidence sur palmier à huile. L'étude de Steyaert (1967), répertoriant les espèces de *Ganoderma* palmicoles, fait état de 6 espèces observées sur le palmier à huile : *G. boninense*, *G. chalceum*, *G. miniatocinctum*, *G. tornatum*, *G. xylonoides*, *G. zonatum*. Toutes ne sont pas considérées comme pathogènes. Trois d'entre-elles sont considérées comme saprophytes : *G. chalceum*, *G. tornatum*, *G. xylonoides*. Jusqu'alors, les études de classification des spécimens récoltés étaient essentiellement basées sur des critères morphologiques basés sur l'observation macroscopique et microscopique des sporophores : forme, structure et couleur de la cuticule, forme et couleur de l'hyménium, forme et structure des spores... Ho et al. (1986) donne une description morphologique du sporophore de *G. boninense* : le sporophore apparaît dans un premier temps comme une structure blanche en forme de bouton et qui évolue sous en forme de dôme au bout de 2 jours. Le sporophore peut être sessile ou posséder un pied dont l'apparence est blanche à sa formation, puis tourne au brun au bout de 14 jours. Concernant le sporophore lui-même, il est formé de cercle concentrique, possède une cuticule de couleur brun-orangé, ainsi qu'une bande de couleur blanche sur le pourtour du sporophore. Sa taille atteint une moyenne de 3,9 cm de longueur et 2,3 cm en épaisseur. L'hyménium est de couleur blanche en surface et les basidiospores sont libérées au bout de la cinquième semaine. Une coupe transversale du sporophore laisse apparaître sous la cuticule une fine couche de tissu jaunâtre (Steyaert 1967).

L'essor de la biologie moléculaire et des techniques de séquençage dans les années 90 a permis d'affiner cette classification et d'identifier de nouvelles espèces notamment grâce

aux marqueurs moléculaires, dont les marqueur du type ITS (Internal Transcribed Spacer) particulièrement utilisés en phylogénie (Baldwin et al. 1995). Les études de Moncalvo (1995 b, 2000) ont apporté quelques clarifications à la classification de ces espèces et à l'attribution d'un échantillon à une espèce donnée. Moncalvo (2000) a montré que la grande majorité des échantillons de sporophore obtenus des plantations de palmier à huile en Asie du Sud-Est appartiennent à l'espèce *G. boninense*. En revanche, les travaux d'Ildris et al. (2001) démontrent la présence récurrente de deux espèces dans les palmeraies de Malaisie péninsulaire : *G. miniatocinctum* et *G. zonatum*, mais n'apporte aucun fait concernant leurs réelles implications dans l'infection du palmier.

#### 2.2.2.2. Dispersion et mode d'infection

La dissection de palmiers infectés par *G. boninense* dès les stades les plus précoces de la maladie a montré que le début de l'infection commence au niveau des racines (Figure 9). Le mycélium, pénètre l'épiderme et l'endoderme et se développe jusqu'au cortex racinaire mais ne colonise pas les couches cellulaires les plus profondes, ni les vaisseaux de circulation de la sève (phloème et xylème) (Rees et al. 2009). Par la suite, la colonisation des tissus

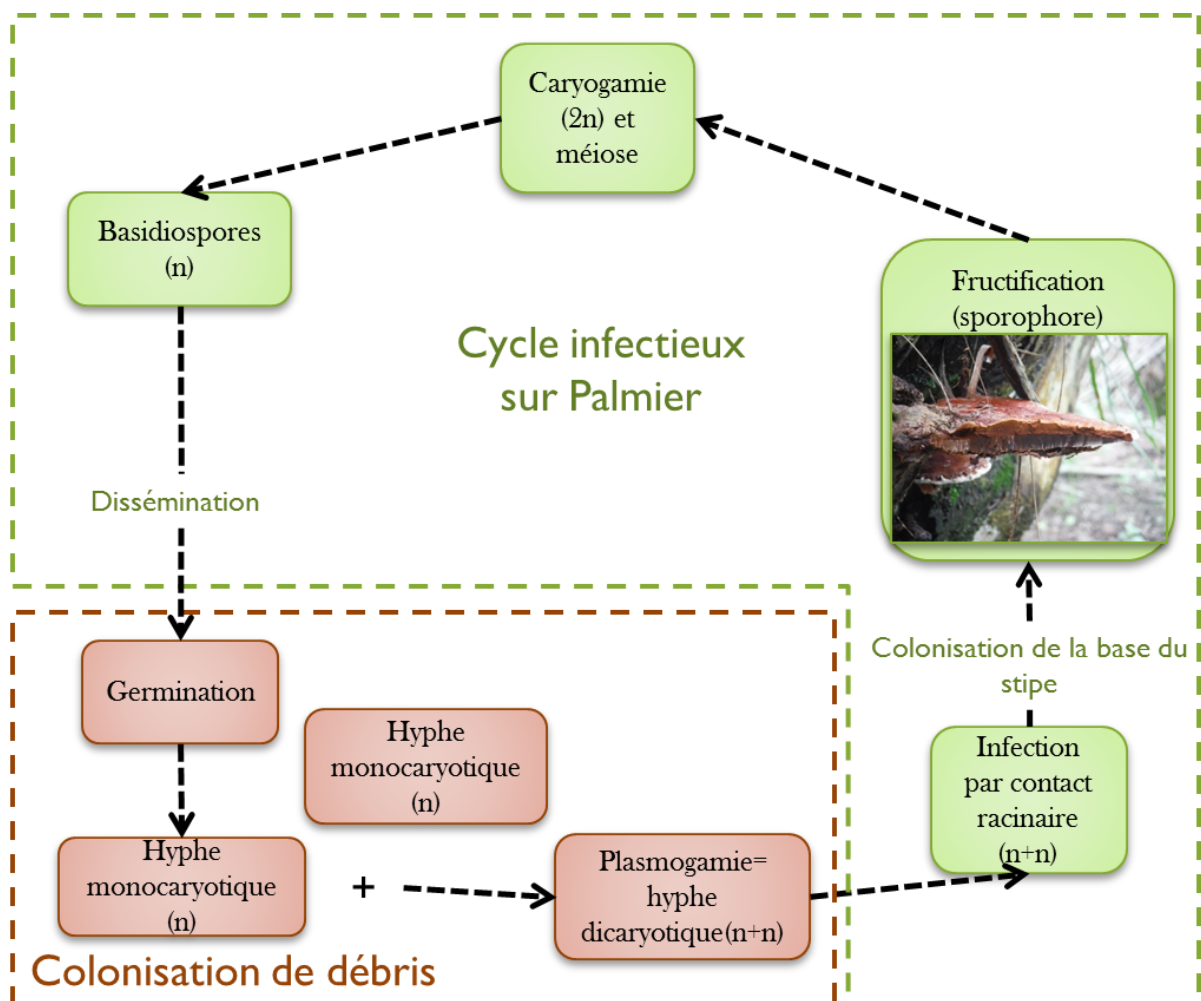


Figure 9. Représentation schématique du cycle de reproduction sexuée de *G. boninense*

progresses le long des racines jusqu'à atteindre les tissus de la base du stipe.

Les premiers travaux concernant l'étude de la propagation et de la dispersion de la BSR (Turner 1965 b) ont montré que les palmiers voisins d'un palmier infecté sont plus fréquemment infectés. Ces observations en conditions naturelles ont conduit à l'hypothèse d'une propagation préférentielle du pathogène entre palmiers par contact racine-racine. En effet, *G. boninense* apparaît être un faible compétiteur dans le sol ne pouvant subsister dans ces conditions que de façon saprophytique sur une masse compacte importante de tissus de palmier (Turner 1981). Cette dernière va devenir alors une source d'inoculum potentiel vis-à-vis des palmiers voisins. Ces résultats sont corrigés par les travaux de Hasan (1998) qui démontrent qu'une source réduite dans le sol de tissu comme des racines infectées peut constituer une source d'inoculum suffisante pour causer de nouvelles infections.

Plus récemment, Pilotti et al. (2002) ont mis en évidence un système de reproduction sexuée multi-allélique et bi-factoriel. Ce système de reproduction est considéré par les auteurs comme un des systèmes les plus complexes découvert à ce jour chez les champignons. Ce dernier se caractérise par un coefficient de restriction de consanguinité de 25% (coefficient correspondant au maximum d'autofécondation) favorisant les croisements intra-population. Les auteurs émettent l'hypothèse que le maintien de la diversité génétique au sein d'une population possédant un tel système, requière un recours important à la reproduction sexuée. Les résultats obtenus grâce à la découverte du système de compatibilité entre isolats de *G. boninense* (Pilotti et al. 2003), ainsi que par marquage moléculaire (Miller et al. 1999, Rees et al. 2012), montrent une répartition aléatoire des foyers d'infection et mettent en évidence une variabilité importante des génotypes présents au sein d'une zone géographique restreinte. Même si plusieurs génotypes ont pu être identifiés au sein d'un même stipe, empêchant d'étudier l'isolat responsable de la primo-infection, une telle diversité des génotypes suggère propagation importante de la maladie par le biais des basidiospores.

Pour conclure sur les modes de propagation, la propagation par contact racinaire ainsi que la dispersion de la maladie par les basidiospores semblent donc toutes deux légitimes mais ne semblent pas intervenir de façon égale.

### 2.2.3. Méthodes de contrôle de la maladie au champ

Le contrôle des maladies au champ repose majoritairement sur la capacité à prévenir ou à prédire l'infection par le pathogène. Cette anticipation est souvent associée à l'établissement de test précoce de détection du pathogène au sein des plantes. Concernant notre pathosystème, le manque de connaissances approfondies sur la biologie du pathogène, sur son mode d'infection et de colonisation des tissus hôtes rend difficile la création d'un test de détection précoce fiable. Des études ont tenté de mettre au point un tel test sur la base de techniques comme la PCR (Idris 2003, Utomo et al. 2005, Idris et al. 2010) ou l'utilisation d'anticorps avec un test ELISA (Idris et al. 2008) mais sans succès. Les

tests sont trop sensibles aux contaminants ou non spécifiques concernant l'espèce de *Ganoderma* à identifier. Les moyens de lutte actuellement utilisés en plantation contre *G. boninense* sont donc pour la plupart des techniques culturales appliquées à d'autres types de cultures pérennes, comme l'attention apportée aux blessures des palmiers provoquées par leurs exploitations ou l'abattage des palmiers avant qu'ils n'atteignent un âge trop avancé, favorisant l'infection par *G. boninense* (Flood et al. 2000 b).

#### 2.2.3.1. Pratiques culturales

Le **buttage** est une technique consistant à ramener de la terre à une hauteur de 75 à 90 cm autour du stipe du palmier dans le but d'initier de nouvelles racines, et de ce fait, prolonger la vie du palmier infecté. L'effet positif de cette pratique est largement discuté entre les planteurs et souvent soumis à controverse.

La **chirurgie** consiste à retirer tous les tissus infectés et dégradés du palmier afin de laisser uniquement les tissus sains. Décrite la première fois par Turner (1981), elle s'accompagne le plus souvent d'un buttage du palmier pour favoriser la formation de nouvelles racines sur la base des tissus retirés. Cette technique s'avère cependant de faible efficacité pour la sauvegarde du palmier.

Ces deux techniques, sont rarement appliquées en plantation industrielles de par la taille des surfaces plantées et donc les moyens humains et le temps nécessaire à leur mise en œuvre. Par ailleurs leurs effets positifs sur le palmier sont controversés et de plus le maintien en survie d'un arbre infecté de quelques mois à 1 an maximum représente une importante source d'inoculum pour les palmiers voisins. Certaines sociétés utilisent ces approches exclusivement dans les jardins semenciers pour essayer de prolonger la vie de palmiers à très forte valeur ajoutée qui sont utilisés dans des programmes d'amélioration.

L'**éradication des palmiers infectés** consiste à abattre les individus atteints de la maladie et d'en extirper le bol racinaire. En général ce dernier est laissé en surface à proximité dans une zone non ombragée. Dans certaines sociétés, ce bol racinaire est déchiqueté au préalable. Flood et ses collaborateurs (2005) ont montré qu'un stipe préalablement colonisé par le pathogène ne représente pas une source d'inoculum pour les palmiers voisins si ce dernier n'est pas enfoui dans le sol. Plusieurs études ont également démontré que *G. boninense* est très sensible à la chaleur à partir de 32°C (Rees et al. 2007). Cette observation tend à confirmer l'effet négatif sur le maintien d'une charge d'inoculum active dans un bol racinaire laissé en surface et au soleil, ainsi que sur l'incapacité du pathogène à se propager dans ces conditions. Il est important de souligner avant les années 1970/1980, en association à cette pratique culturale, les tissus étaient brûlés, ainsi que l'intérieur du trou formé par l'extirpation du bol racinaire. Ce brûlage est aujourd'hui interdit en Asie du Sud-Est de par l'importante pollution qu'il engendrait. Cependant dans certaines régions d'Afrique, cette méthode reste encore utilisée.



La **jachère** a également montré un effet significatif sur le taux d'infection des palmiers replantés. Cependant des études complémentaires seront nécessaires afin de déterminer la durée de jachère et le type de plante pouvant être employé pour pallier à la baisse des revenus causée par cette jachère.

Des tranchées de 2 mètres de largeur sur 2 mètres de profondeur autour du palmier infecté sont également utilisées en de rares occasions, mais permettent un ralentissement significatif de la propagation de la maladie (Hasan 1998). Cependant, cette méthode est difficilement envisageable dans les plantations industrielles de plusieurs milliers d'hectares.

#### 2.2.3.2. Lutte chimique vs lutte biologique

Quelques études ont tenté de tester l'efficacité de traitement fongicide par injection dans le palmier malade (Ariffin 1993, George 1996, Idris et al. 2004, Idris et al. 2010). Des fongicides comme Dazomet ou l'hexaconazole n'ont pour l'instant révélé aucune efficacité dans le contrôle de la maladie (Chung 2011). Cependant, une prolongation de la vie de l'arbre a pu être constatée (Idris 2012).

En parallèle des tests de lutte chimique, plusieurs études concernant un contrôle biologique de la BSR ont été lancées à partir des idées développées par Varghese et al. (1976). Si les études de Sariah et al. (2000) et Yow et al. (2001) basées sur l'utilisation de champignon du genre *Trichoderma*, connus pour leur effet antagoniste sur d'autres champignons (Dennis et al. 1971, Tronsmo et al. 1977), donnent des résultats souvent controversés pour la lutte contre *Ganoderma boninense*, Susanto et al. (2005) montrent une réduction significative des taux d'infection dans les essais employant le champignon *Trichoderma harzianum* et la bactérie *Gliocladium viride*. Ces études ayant été menée en milieu contrôlé et sur des fonds génétiques de palmiers différents, il n'est pas possible de statuer sur l'efficacité réelle de ces antagonistes de *G. boninense* en plantation.

#### 2.2.3.3. Sélection variétale

La sélection pour le développement de matériel résistant commence avec l'observation en Indonésie de différence dans l'incidence de la BSR entre les palmiers issus de fonds génétiques d'Afrique de l'ouest et ceux issus de matériel Deli (Akbar et al. 1971). Par la suite, De Franqueville et al. (2001) met en évidence une différence significative dans l'incidence de *G. boninense* entre les familles de 8 essais de sélection sur 12 des essais mis en place, ainsi qu'entre clones sur 6 des 7 essais de clones mis en place. Ces différences d'incidence varient de 12 à 75% en fonction de la famille considérée et ont été mesurée après une durée de culture de 24 ans. Actuellement, la sélection de palmiers résistants, ou plutôt tolérants, passe par des tests précoces d'inoculation de *G. boninense* sur graines germées en préépinière (Figure 10). La reproductibilité des mesures de tolérance effectuées au champ et en pépinière suggère que les mécanismes de réponse à l'infection par *G. boninense* restent identiques entre le stade juvénile du palmier et le stade adulte après mise en terre en plantation et que les conditions environnementales (apport en intrant, température,

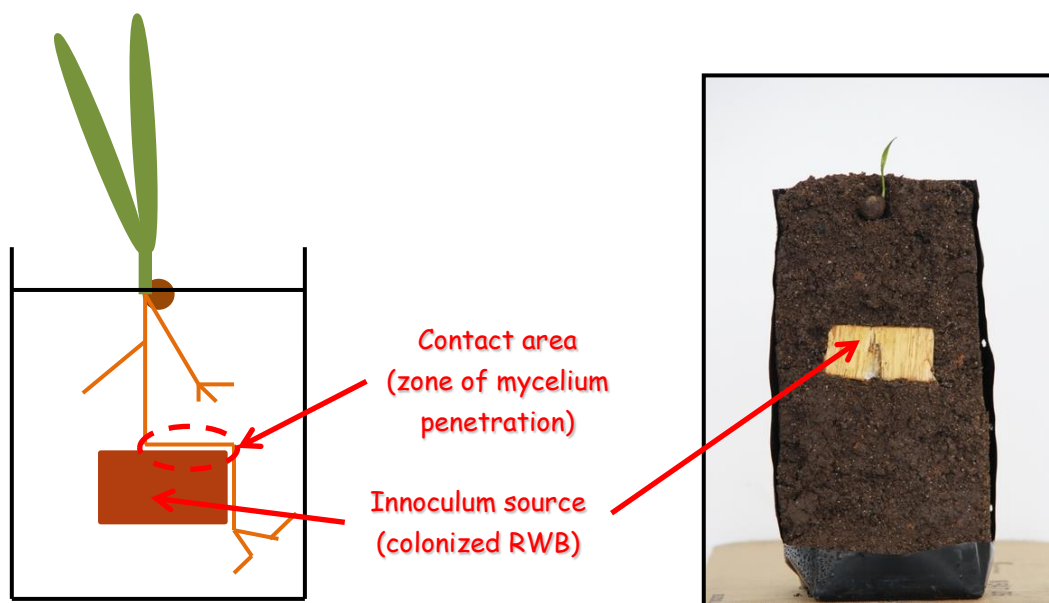


Figure 10. Schéma et photo du dispositif d'inoculation de la graines germées au stade 2 feuilles en préépinière

humidité, taux d'ensoleillement, ...) n'influent que marginalement sur la capacité de résistance du palmier. La sélection des individus tolérants est basée sur le taux d'incidence de la maladie sur le matériel testé, donc sur le nombre d'individus du même matériel sur lesquels des symptômes ont pu être observés pour une durée identique à tous les tests (Breton et al. 2006).

Dans le but d'améliorer les programmes de sélection, il est nécessaire de mieux comprendre la biologie de l'agent pathogène et particulièrement les interactions hôtes-pathogènes mises en place au cours de l'infection. Au cours de ces dernières années quelques études ont pu mettre en évidence des variations dans l'expression de gènes potentiellement impliqués dans les mécanismes de défense du palmier, à l'aide d'études transcriptomiques de palmiers infectés par *Ganoderma* (Tan et al. 2013, Tee et al. 2013, Ho et al. 2014). Ces gènes candidat nécessiteront une validation de leurs implications dans les mécanismes de réponse à l'infection par *Ganoderma* ainsi que de leurs rôles dans la résistance (ou tolérance) du palmier à la BSR. Par la suite, le suivi de l'expression de ces gènes, s'ils sont bien spécifiques à la BSR, pourrait être un marqueur intéressant pour la détection précoce d'une infection par *G. boninense*.

### 3. Les objectifs de la thèse

Le travail de cette thèse est orienté vers l'étude de la biologie et particulièrement de la biologie moléculaire de l'espèce *G. boninense* afin de mieux comprendre les facteurs influençant sa diversité et sa dispersion, ainsi que les mécanismes qui régissent son interaction avec le palmier à huile. Dans cette optique nous avons mêlé des approches de génomique, de génétique des populations et de transcriptomique. Un effort particulier d'échantillonnage à plusieurs échelles géographiques a été concentré sur les zones de

Malaisie péninsulaire et de l'île Indonésienne de Sumatra qui représentent les zones historiques de la culture industrielle du palmier à huile. Cette zone géographique constitue un environnement idéal pour l'étude de l'histoire démographique des populations de *G. boninense*.

Dans un premier temps, le travail a consisté à séquencer le génome de *G. boninense* dans le but de développer un ensemble d'outils moléculaires, nécessaires pour l'étude précise de la structure génétique des populations de *G. boninense* en Asie du Sud-Est et de l'histoire démographique de ces populations. Un jeu de marqueurs microsatellites a donc été développé sur la base des données génomiques afin de répondre à une première question : combien de groupe génétique distinct existe-t-il au sein des plantations d'Asie du Sud-Est ?

Dans un deuxième temps, nous avons souhaité étudier de manière précise la structuration du groupe d'échantillons provenant de nos récoltes dans la zone historique de culture industrielle du palmier à huile. L'étude restreinte à ce groupe d'individus est destinée à nous renseigner sur différentes caractéristiques biologiques de *G. boninense* en plantation industrielle : Quelle est le mode et la capacité de dispersion de *G. boninense* ? Existe-t-il une spécialisation des génotypes en fonction des plantations ou des blocs liés aux génotypes de l'hôte, au nombre de génération... ? Y'a-t-il eu un saut d'hôte à l'origine de l'adaptation de *G. boninense* au palmier à huile ?

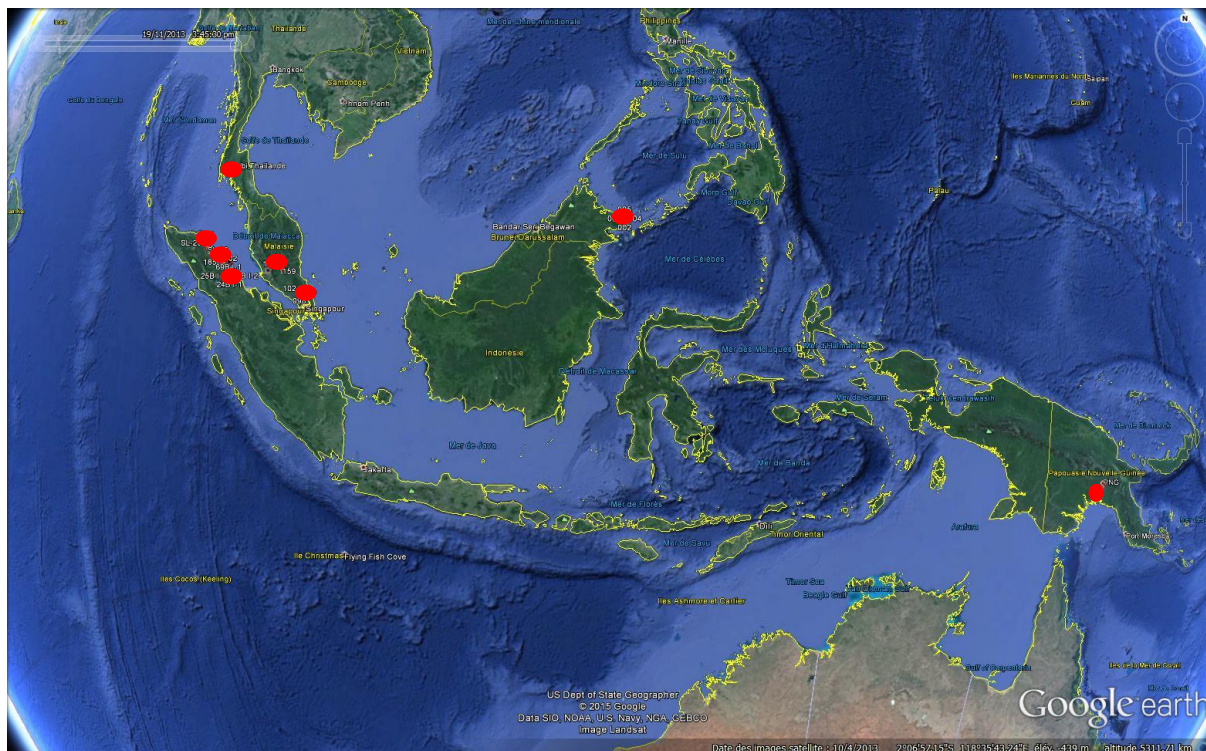
De plus, nous avons également voulu connaître à l'aide de ces marqueurs microsatellites la distance génétique entre les individus d'Asie du Sud-Est et ceux d'Afrique.

Enfin, dans un dernier temps, nous avons voulu explorer les bases génétiques de la variation d'agressivité observée par Breton et al. (2006) entre les différents isolats. Notre problématique était donc d'observer si des variations génétiques d'expression étaient à l'origine de ces variations d'agressivité entre isolats : cette variation est-elle due à la présence de gènes exprimés spécifiquement dans les isolats agressifs ? Ou cette variation est-elle due à variation quantitative d'expression de gènes ? Existe-t-il une expression spécifique de gènes liés à la reconnaissance de l'hôte par le pathogène ?

Les données de transcriptomique obtenues de cette expérimentation permettront également de produire une annotation complète du génome de *G. boninense* obtenus lors de la première partie du travail.

## **CHAPITRE II :**

# **MATERIELS ET METHODES**



**Figure 11. Carte de la région d'Asie du Sud-Est regroupant l'intégralité des zones d'échantillonnages. Les cartes représentant chaque plantation sont disponibles en annexe 1**

Dans ce chapitre nous présenterons la méthode d'échantillonnage mise en œuvre pour répondre à la problématique concernant la diversité génétique du pathogène au champ, ainsi que les différentes étapes ayant permis le développement des marqueurs moléculaires. La méthodologie ainsi que les techniques employées dans le but de répondre aux objectifs de l'étude des bases moléculaires de l'agressivité de *G. boninense* seront présentées dans un chapitre indépendant.

## **1. Echantillonnage des isolats de *G. boninense***

Au cours de cette thèse deux missions d'échantillonnage ont été réalisées en palmeraie chez nos partenaires en Malaisie (FELDA) et en Indonésie (PT SOCFINDO) au cours de la première année (Figure 11). Ces deux missions ont permis la récolte d'un total de 401 échantillons répartis sur 9 plantations

Concernant la récolte indonésienne, 6 plantations de Sumatra ont été échantillonnées constituant une collection de 275 individus. Ces échantillons sont répartis au sein de plusieurs blocs sur chaque plantation :

- Bangun Bandar : 107 échantillons pour 13 blocs,
- Mata Pao : 62 échantillons pour 9 blocs,
- Tanah Gambus : 64 échantillons pour 10 blocs,
- Sei Liput : 27 échantillons sur 12 blocs,
- Aek Loba : 8 échantillons provenant chacun d'un bloc différent,
- Plantation du sud de Sumatra : 7 échantillons provenant également de différents blocs.





Figure 12. Echelle de symptômes des arbres atteint de la BSR. A) aucun symptômes apparents. B) Symptômes légers. C) Symptômes importants. D) Symptômes importants

La récolte en Malaisie a été réalisée sur 3 plantations et comptabilise un total de 126 échantillons. Ces échantillons ont été obtenus dans des plantations plus distantes qu'à Sumatra, ayant pour conséquence un échantillonnage moins important :

-Malaisie péninsulaire : Negeri Sembilan représentant un total de 47 échantillons au travers 2 blocs, Johor représentant 45 échantillons répartis sur 3 blocs,

- Bornéo : plantation de l'Etat du Sabah représentant un échantillonnage de 34 sporophores répartis sur 7 blocs. Les échantillons récoltés dans cette plantation ont été récoltés en amont de la mission par notre partenaire du FELDA et conditionnés au cours de la mission.

Au cours de ces missions, plusieurs informations concernant les échantillons ont été relevées au cours de l'échantillonnage : les coordonnées GPS de chaque sporophore, le nom de la plantation, le numéro de bloc, la génération de plantation, l'année de plantation, l'état général du palmier (vivant, mort encore debout ou mort et tombé) ainsi que le niveau de symptômes de l'arbre (Figure 12). Pour certaines plantations de Sumatra, le fonds génétique de l'arbre, sur lequel a été récolté le sporophore, a pu être renseigné (Tableau récapitulatif en annexe).

L'échantillonnage obtenu à la suite de ces missions a été complété par la réception de sporophores collectés par nos partenaires et provenant d'Afrique (Cameroun, 32 échantillons), de Thaïlande (4 échantillons), ainsi que de Papouasie-Nouvelle-Guinée (11 échantillons). Ces différents niveaux d'échantillonnage nous ont permis d'étudier la diversité du pathogène à des échelles géographiques plus ou moins larges. Tous les sporophores ont été disséqués afin d'enlever la majeure partie des contaminants de surface et lyophilisés afin de prévenir la dégradation du matériel génétique, dans le but d'un stockage à long terme.

## **2. Séquençage du génome**

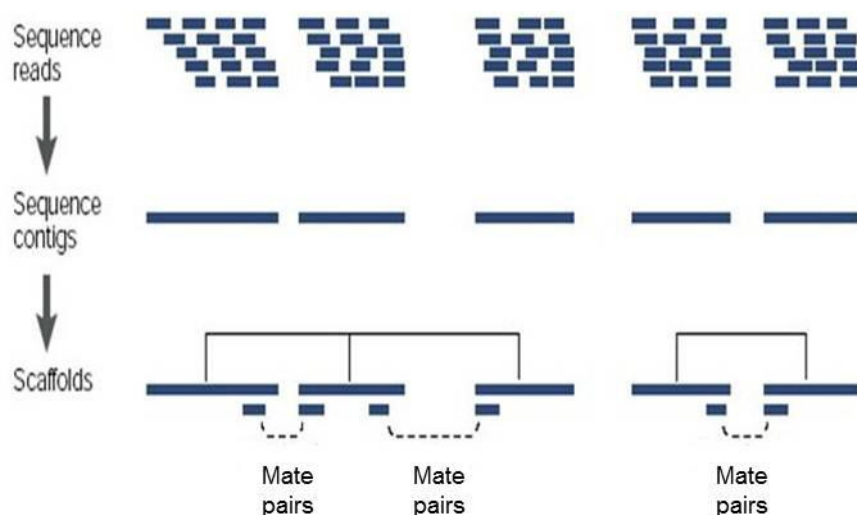
La séquence référence du génome de *G. boninense* n'est jusqu'à présent pas disponible dans la littérature et les bases de données mondiales. Le projet dans lequel s'intègre cette thèse a donc inclus la production d'un génome de référence de *G. boninense*. Cette séquence de référence a permis le développement de nouveaux marqueurs moléculaires ainsi que l'obtention d'une séquence d'ADN génomique de référence pouvant aider à l'exploitation des données de transcriptomique qui doivent permettre de mieux comprendre les bases génétiques de l'agressivité. La production de cette séquence de référence a nécessité le séquençage d'une souche mycélienne pure de *G. boninense* considérée comme référente. La souche choisie fut la souche portant le code NJ3, considérée comme d'importance prioritaire par notre partenaire PalmElite car utilisée comme référence dans les tests précoces de résistance en pré-pépinière.. Les cultures pures de *G. boninense* sont obtenues par dissection de la partie interne du sporophore, stérilisation en surface à l'éthanol 70% de petits cubes de tissus qui sont par la suite déposés en boîte de pétri sur milieu pauvre en nutriments WA (Water-Agar) en présence d'antibiotiques (ici streptomycine et chloramphenicol) et entre 27 et 29°C. Ce milieu permet la croissance d'un mycélium fin que l'on transfère sur un milieu nutritif PDA (Potatoe Dextrose Agar).



## 2.1. Optimisation du protocole d'extraction d'ADN

Plusieurs protocoles d'extraction ont été testés sur le mycélium : le kit MasterBeads de la société Ademtech, le protocole d'extraction d'ADN Dolezel-MATAB (1989), et modification du protocole Dolezel contenant une première étape de digestion enzymatique suivie d'une extraction MATAB classique. Les résultats les plus probants ont été obtenus par l'extraction MATAB précédé de la digestion enzymatique (Annexe 2). En effet, le cocktail enzymatique GLUCANEX commercialisé par la société SIGMA-ALDRICH est composé d'un ensemble de  $\beta$ -glucanase, cellulase, protéase et chitinase permettant une digestion de la paroi épaisse de *G. boninense*. Cette première digestion de la paroi permet de s'affranchir d'une partie des polysaccharides contenue dans la cellule fongique et permet un accès facilité au contenu cellulaire. Les polysaccharides sont l'un des polluants majeurs des échantillons d'ADN fongiques après extraction les rendant impropre à un séquençage de qualité. L'élimination des polysaccharides en pré-extraction est donc une étape importante du processus d'extraction. La charge positive des polysaccharides est identique à celle de la molécule d'ADN et rend difficile leur séparation une fois dans le milieu post-extraction.

## 2.2. Choix du protocole de séquençage



**Figure 13.** Schéma générale des étapes d'assemblage de génome ou transcriptome

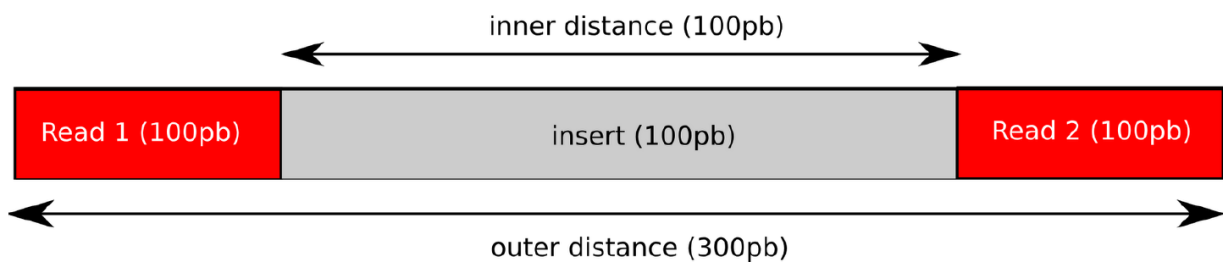
Afin de déterminer le protocole de séquençage adapté à la problématique envisagée, il est important de connaître quelques notions d'assemblage de génomes à partir de données de séquençage.

La figure 13 montre le processus « classique » d'assemblage. Il est nécessaire de rappeler que les techniques de séquençages disponibles au début de la thèse ne peuvent séquencer des fragments d'ADN plus long que 500 bp pour la technologie Roche 454 et 150 pour la technologie CRT Solexa/Illumina. L'ADN génomique doit donc être découpé avant le séquençage.

Le processus :



- 1- L'ADN lors du séquençage est fragmenté en plusieurs fragments de longueur égale, la plupart du temps par une fragmentation mécanique.
- 2- Des adaptateurs sont ajoutés à l'une des extrémités (technique de séquençage **single-end**) ou aux deux extrémités des fragments d'ADN (technique de séquençage **paired-end**). Les adaptateurs sont des séquences courtes d'ADN permettant l'utilisation d'une seule et même amorce pour l'amplification de l'ensemble des fragments. Ces deux premiers points forment ce que l'on appelle en séquençage la préparation de la **banque** (librairie en anglais).
- 3- La librairie obtenue est transférée sur le support de séquençage où les fragments double brin sont deshybridés et séquencés individuellement.
- 4- Les séquences obtenues après acquisitions des données sont appelées **reads**. Ces reads sont des fragments d'ADN amplifié et séquencé aléatoirement ce qui occasionne des chevauchements (ou de la redondance) entre les différents reads.
- 5- Ces chevauchements sont alors utilisés dans le but d'obtenir les séquences consensus de l'assemblage de tous les reads. Ces séquences sont appelées **contigs**.



**Figure 14.** Fragment d'ADN obtenu après préparation de la librairie de séquençage

La différence entre les séquences obtenues à partir de protocole single-end ou paired-end réside dans le fait que l'ajout d'adaptateurs aux deux extrémités des fragments générés dans le protocole paired-end permet le séquençage de reads par paires et donc l'obtention de séquence de fragments plus longs (2 x 500 bp pour la technologie Roche 454 et 2 x 150 bp pour la technologie illumina). Les fragments d'ADN obtenus par ce protocole ont la forme présentée sur la figure 14. Ils présentent une distance « inner » et une distance « outer ». La distance « inner » est la distance qui sépare les deux reads, sans inclure la taille des reads dans le calcul de la distance. La distance « outer » inclue la taille des reads dans le calcul de la distance. Si la distance « outer » est égale à la taille des reads (par exemple une longueur de 100 bp comme la longueur de l'insert) alors la paire de reads sera en chevauchement total. Le protocole paired-end permet donc l'obtention de reads pour des fragments d'ADN beaucoup plus long et améliore le processus d'assemblage des contigs.

Une adaptation du protocole de préparation des banques paired-end a permis l'utilisation d'inserts plus grands que les 100 bp habituellement utilisées. Ce protocole est appelé **mate pair** et repose sur le principe de circularisation des fragments d'ADN pouvant intégrer des inserts jusqu'à 6 Kb (Kilobase) pour les technologies Illumina.

Ces reads pairés d'environ 6 Kb permettront un assemblage plus complet grâce à une étape de plus dans le processus d'assemblage appelé **scaffolding**. Le scaffolding consiste à rassembler les reads pairés (reads séparés par un insert) qui ont été assemblés en contigs afin de rassembler plusieurs contigs formant ainsi des scaffolds.

Si l'utilisation de la technologie Roche 454 paraît être la plus intéressante en matière de longueur de séquences obtenues, elle offre une faible profondeur de lecture (nombre de reads générés correspondant à la même séquence) comparé à la technologie Illumina. Une faible profondeur de lecture peut entraîner l'intégration de reads mal amplifiés (dérapage de la taq polymérase, incorporation de mauvais nucléotides) et donc générer des erreurs de séquence, qui sont corrigées grâce à l'amplification d'une quantité importante de reads sur une même séquence.

Notre projet de séquençage a donc utilisé un mélange des deux technologies en utilisant à la fois du séquençage 454 en single-end pour la génération de séquences longues et également la technologie de séquençage illumina en paired-end avec un insert de 3 kb pour obtenir une profondeur de lecture accrue évitant ainsi le maximum d'erreurs de lecture. Un nombre de 58 559 370 reads a ainsi été généré (2,5 millions de reads 454 d'une longueur moyenne de 379 bp et 56 millions de reads Illumina d'une longueur moyenne de 51 bp) et un total de 21 425 contigs ont été assemblés. Le processus de scaffolding a ainsi assemblé les 21 425 contigs à l'aide des reads pairés de la technologie Illumina et a permis de former 5 664 scaffolds. À partir de cet assemblage la taille du génome de *G. boninense* a été estimée à 70 Mb (Megabase). L'assemblage a ensuite permis le développement des marqueurs microsatellites nécessaires à l'étude de la diversité.

### **3. Développement des marqueurs moléculaires**

Cette partie du chapitre II traite le développement des marqueurs microsatellites à partir d'un assemblage automatisé du génome de *G. boninense*. Un sous-échantillonnage d'individu provenant des missions de collecte, ainsi que deux isolats d'espèce proche de *G. boninense* issus de la mycothèque du CBS-KNAW ont été utilisés dans le but (i) de tester l'efficacité des marqueurs microsatellites sur un matériel différent que la culture mycélienne ; (ii) d'obtenir une première idée de la structuration des populations au sein des plantations de nos partenaires et (iii) de statuer sur le potentiel de transfert interspécifique de ces marqueurs. Cette article sera publié en 2015 dans la revue *Mycological progress* et a été accepté.

## Identification and development of new polymorphic microsatellite markers using genome assembly for *Ganoderma boninense*, causal agent of oil palm basal stem rot disease

Mercière Maxime<sup>(1)</sup>, Laybats Anthony<sup>(1)</sup>, Carasco-Lacombe Catherine<sup>(1)</sup>, Tan Joon Sheong<sup>(2)</sup>, Klopp Christophe<sup>(4)</sup>, Tristan Durand-Gasselin<sup>(3)</sup>, Sharifah Shahrul Rabiah Syed Alwee<sup>(2)</sup>, Létizia Camus-Kulandaivelu<sup>(1)</sup>, Breton Frédéric<sup>(1)</sup>

CIRAD (1), FELDA (2), PalmElit (3), INRA (4)

(1) CIRAD, UMR 108 - AGAP Avenue Agropolis, 34398 Montpellier, Cedex 5 France. Email: maxime.merciere@cirad.fr

(2) FELDA GLOBAL VENTURES RESEARCH & DEVELOPMENT SDN BHD.  
Felda Biotechnology Centre, PT 23417, Lengkuu Technology, 71760 Bandar Enstek, Negeri Sembilan, Malaysia.

(3) PalmElit SAS, Bat 14, Parc Agropolis, 2214 Bd de la Lironde, 34980 Montferrier sur Lez, France.

(4) INRA, plate-forme Bioinfo Genotoul, UR875, Toulouse, CS 52627, 31326 Castanet-Tolosan, France.

### Abstract

*Ganoderma boninense* is a telluric lignicolous basidiomycete and the causal agent of basal stem rot, one of the most devastating diseases of the oil palm (*Elaeis guineensis*). While the fight against *G. boninense* is of major concern in Southeast Asia, little information is available about the genetic diversity and evolutionary history of the fungus. In this context, the development of an informative molecular marker set to characterize *G. boninense* diversity is a key step towards understanding the biology of this pathogen. A *G. boninense* draft genome sequence of 63 Mbp, assembled using 454 and Illumina sequencing technologies, was used to identify and develop a set of microsatellite markers (simple sequence repeats - SSRs). A total of 2,487 SSRs were identified, for which 145 SSR primer pairs were designed. These SSRs are characterized by di- to hexa-nucleotide motifs with 5 to 34 repetitions. Ninety-seven SSR loci were successfully amplified on an initial small set of *G. boninense* isolates from Indonesia. A collection of 107 isolates from several regions in Southeast Asia was screened to characterize each locus for allele number, polymorphism information criterion and the presence or absence of null alleles at each locus. These results allowed us to propose an effective set of 17 SSRs for studying *G. boninense* genetic diversity.

**Keywords:** *Ganoderma boninense*, oil palm, basal stem rot, microsatellite, genome sequence

### Introduction

The availability of new molecular biology tools along with the exploration of genomic information has made it possible to develop many genetic markers including microsatellites. Microsatellites, also known as Simple-Sequence Repeats (SSR), are DNA regions comprising short tandem repeat units from 1 to 6 base pairs. Polymorphic microsatellites have become one of the most popular and commonly used molecular markers for studying genetic diversity, genealogy and molecular ecology (Jarne et al. 1996). In particular, microsatellites are still the most powerful and most commonly used tool for studying population genetics and the geographical spread of plant pathogens (Schoebel et al. 2013, Presti et al. 2014). This success results from their codominance, high polymorphism, high allele number and relatively high abundance in most genomes and the ease of typing by polymerase chain reaction (PCR).

Until recently, developing microsatellite markers for a previously unsequenced species involved the construction of an enriched DNA library by cloning sequences in plasmids and sequencing them using the Sanger method

(Zane et al. 2002), which was expensive and time-consuming. The development of Next Generation Sequencing (NGS) technologies makes available large amounts of genomic sequence at a low cost (0.05 \$ for 1 Mbp in April 2014 against 5292 \$ for 1 Mbp in September 2001; NIH estimate: <http://www.genome.gov/sequencingcosts>) and provides a faster and cheaper solution for finding microsatellites without *a priori* knowledge of the target species. In recent years, NGS platforms such as Illumina GAII or Roche 454 have made larger sets of microsatellite markers available by exploring complete genome sequences, even for non-model species (Abdelkrim et al. 2009, Zalapa et al. 2012, Cai et al. 2013, Schoebel et al. 2013). Until now, most studies searched raw NGS data for microsatellite patterns to design primers (Gardner et al. 2011). Using NGS raw data, i.e. raw reads, has a number of limitations. The size of raw reads from the 454 (500 bp) and Illumina (250 bp) systems is often too short to design primers for both sides of the SSR, which is a critical step in developing microsatellite markers (Castoe et al. 2010). Proper primer design for SSR genotyping should, therefore, satisfy several conditions including a minimum PCR product size of 100 bp to avoid confounding artefacts caused by primer peak or same melting temperature ( $T_m$ ) for both forward and reverse primers. Castoe et al. (2010) were able to design primers for only 49.6% of the SSRs they identified in their 454 raw read data set and Cai et al. (2013) were able to design primers for only 32.7% to 40.1% of the SSRs they identified in their Illumina raw read data set. Moreover, sequence redundancy in raw reads is very challenging when designing locus specific primers because determining whether reads with similar sequences correspond to a single locus or to duplicated regions is non-trivial (Castoe et al. 2010). New approaches based on the exploration of genome assemblies seem to be a promising way of circumventing both problems caused by the use of raw reads (Cai et al. 2013). Read assembly consists in aligning and merging reads into much longer DNA/RNA fragments in order to reconstruct the original molecule. Read assembly thus provides much longer sequences of about 1 to 10 kbp for a raw assembly against a maximum of 600 bp for 454 raw reads or 300 bp for Illumina raw reads (Zalapa et al. 2012, Wei et al. 2014) and considerably reduces data complexity.

Despite the advent of NGS, developing fungal microsatellite markers is still challenging (Dutech et al. 2007, Schoebel et al. 2013). While the small size of fungal genomes, most ranging from 10 to 60 Mbp (Gregory et al. 2007), makes it easier to explore the genome to find microsatellite loci, it appears that microsatellites are less abundant in fungi than in any other organisms (Dutech et al. 2007). Fungal microsatellite motifs are also known to have low repetition numbers (mostly below 8 repetitions) with a high proportion of mononucleotide motifs, making it difficult to find highly polymorphic SSRs (Dutech et al. 2007). Finally, the cross-species transferability of SSRs in fungi seems to be lower than in other organisms (Barbara et al. 2007, Dutech et al. 2007).

In this study, we developed an SSR marker set from a draft genome assembly of *Ganoderma boninense*, a telluric, white rot basidiomycete, which is the causal agent of oil palm basal stem rot (BSR). *G. boninense* is a non-model plant pathogen fungus with a poorly known biology (Ho et al. 1986, Pilotti et al. 2002), but it belongs to the same genus as *Ganoderma lucidum*, a very popular medicinal fungus that has already been sequenced (Chen et al. 2012). *G. boninense* can affect up to 50% of the palm trees in a plantation with impacts ranging from lower productivity to tree death (Corley et al. 2003) and is of major economic importance in Southeast Asia (Turner 1981). As palm oil has the largest share of the world vegetable oil market with 54 million tons produced in 2012 (United States Department of Agriculture: <http://apps.fas.usda.gov/>) and 80% is used by the food industry, BSR may have a serious effect on global food security. Most studies of *G. boninense* were carried out to select resistant oil palm progenies (Idris et al. 2004, Durand-Gasselin et al. 2005, Breton et al. 2006) or to develop management practices to limit the spread of the disease (Susanto et al. 2005). To fight *G. boninense* more effectively we must improve our understanding of its biology. Our knowledge of *G. boninense* physiology (Pilotti 2005, Rees et al. 2007) and genetics (Miller et al. 1999, Pilotti et al. 2003) is insufficient to explain its current range across an area as large as Southeast Asia. In recent years, population genetics has proved to be an important tool for understanding the biology and spread of plant pathogens (Jarne et al. 1996, Travadon et al. 2012 a, Travadon et al. 2012 b). Genetic diversity patterns can provide information about the respective contributions of clonal and sexual reproduction (Travadon et al. 2012 b) and the existence of gene flow barriers (Ali et al. 2014). Neutral molecular markers can be used for reconstructing the main dispersion paths of cosmopolitan or invasive plant pathogens (Montarry et al. 2010) and even lead to the identification of the original hotspot of an emerging disease (Dutech et al. 2012).

This study contributes to our knowledge of *G. boninense* genetic diversity. In this study, we first used a “Seq-Assembly-SSR” approach (Cai et al. 2013) to our preliminary genome assembly based on two sequencing methods (454 and Illumina) to develop a suitable set of SSR markers that was then used to produce an initial description of the genetic diversity of a collection of *G. boninense* isolates from southern Asia. As the microsatellite compatibility between closely related fungal species is often described as anecdotal (Dutech et al.

**Table 1.** Geolocation of the 107 *Ganoderma boninense* samples from the genotyping population

Country	Region	Plantation	Plantation GPS coordinates	Samples
Indonesia	Sumatra	Aek loba	2°37'46.00"N 99°37'42.00"E	AL1 / AL2 / AL3 / AL4 / AL5 / AL7 / AL8
		Mata pau	3°32'52.67"N 99°51'1.17"E	311 / 312 / 318 / 319 / 325 / 334 / 341 / 347 / 357 / 366
		PSBB	3°19'5.32"N 99°2'37.19"E	200 / 210 / 218 / 226 / 261 / 271 / 282 / 290 / 299 / 306
		Sei liput	4°20'32.80"N 98°3'52.50"E	SL01 / SL03 / SL05 / SL08 / SL10 / SL12 / SL17 / SL20 / SL22 / SL24
		Tanah gambus	3°14'26.70"N 99°26'0.17"E	374 / 382 / 389 / 396 / 398 / 406 / 419 / 412 / 420 / 427
		Sabah	5°10'52.41"N 119°12'38.20"E	1 / 2 / 3 / 4 / 5 / 6 / 7 / 10 / 11 / 12 / 22 / 23 / 25 / 26 / 27 / 28 / 30 / 31 / 32 / 33
Malaysia	Borneo			
	Peninsular Malaysia	Johor bahru	1°44'34.27"N 103°38'52.07"E	70 / 71 / 72 / 73 / 74 / 75 / 76 / 77 / 80 / 81 / 82 / 83 / 84 / 85 / 86 / 87 / 88 / 89 / 90 / 91
		Negeri sembilan	2°57'16.08"N 102°29'20.58"E	113 / 114 / 115 / 116 / 117 / 118 / 119 / 120 / 121 / 122 / 123 / 124 / 125 / 126 / 127 / 128 / 129 / 130 / 131 / 133



2007), we also tested the transferability of those markers to a closely related species, *G. resinaceum*. The development of this microsatellite marker set would seem to be essential for gaining a better understanding of diversity of *G. boninense* and may be a first step towards improving crop and disease management practices.

## Materials and methods

### Fungal material

As *G. boninense* is the main species found on oil palms (How et al. 1985), we based our sampling on the morphological traits of basidiocarps as described by Ho and Nawawi (1986) and Flood et al. (2000). The basidiocarps may be stalked or sessile, flat and bracket-shaped. The dorsal surface is glossy, blackish-brown in color with concentric markings. The edge is white when fresh and the under-surface is also white. Five *G. boninense* specimens were collected from infected palms in industrial plantations owned by PT Socfin Indonesia in North Sumatra (Indonesia). Mycelium was purified and maintained in a standard PDA medium (39 g.L<sup>-1</sup> PDA from SIGMA-ALDRICH®, 0.5 g.L<sup>-1</sup> Chloramphenicol, pH 7.0) in the dark at 34 °C. We assumed that these isolates are constituted of dikaryotic mycelium. One isolate, NJ3, which was also used for genome sequencing, and four other isolates were used for the first SSR validation step. To assess the trans-specific amplification pattern for the SSRs, we obtained two isolates of *G. resinaceum*, a species closely related to *G. boninense*, from the CBS-KNAW Fungal Biodiversity Center's culture collection (<http://www.cbs.knaw.nl/index.php/collection>). One hundred and seven fruiting bodies were collected from peninsular Malaysia (40 fruiting bodies), Borneo (20 fruiting bodies) and North Sumatra (47 fruiting bodies). Geographical locations and sample distribution are summarized in Table 1. Each sample was isolated from a single oil palm tree and identified as *G. boninense* using morphological criteria. Only dikaryotic tissues from the fruiting bodies were kept and freeze dried before being packed with silica gel.

### DNA extraction

Dikaryotic tissues from the fruiting bodies and 2-week-old pure cultures were crushed to powder and the DNA was extracted with the MATAB protocol adapted from the CTAB protocol (Doyle 1987) with some modifications to improve the DNA extraction yield (Tris HCl 1M pH 8, NaCl 5M, EDTA 500mM, MATAB, PEG 6000 and sodium sulfite). A first step with cell wall digestion by GLUCANEX (Lysing enzymes from *Trichoderma harzianum*, SIGMA-ALDRICH® buffer (GLUCANEX® 0.5g, NaCl 1M 70ml, H<sub>2</sub>O 30ml, buffer at pH 6) gave a higher DNA yield and removed most of the polysaccharides. The MATAB protocol was also modified for 2 ml tubes to facilitate the extraction of a large set of samples. The DNA concentration (ng.µl<sup>-1</sup>) of each sample was determined by fluorometry (Qubit® 2.0 Fluorometre, Invitrogen) to avoid measurement bias due to the high polysaccharide content.

### Genome sequencing and assembly

The draft genome sequence was obtained using total genomic DNA extracted from the NJ3 isolate. A combination of 454-GS-FLX reads (single reads with theoretical length of 400 bp) and Illumina HiSeq 2000 reads (paired reads with insert of 3 Kb and theoretical length of 50 bp) were used for contig assembly. The reads were assembled using CLC Genomics Workbench 4.6.1 (CLC bio) with standard parameters. The N25, N50 and N75 genome assembly statistics were calculated. N values are standard quality measurement values for genomic data (N25, N50, N75 are calculated by summarizing the lengths of the biggest contigs until you reach N % of the total contig length and the minimum contig length in this set is the number that is usually used to report the N value of a de novo assembly). This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LFMK00000000. The version described in this paper is version LFMK01000000.

### SSR development and primer design

The draft genome of *G. boninense* (NJ3 isolate) was screened with MSATCOMMANDER 1.0.8 (Faircloth 2008) to identify SSRs. The parameters were set to search for di- to hexa-nucleotide SSR perfect motifs with at least 5 repeats. Only candidate SSRs with flanking sequences of at least 400 bp were kept in order to increase the success of the primer design and to increase flexibility in the choice of fragment length. Primers were designed using Primer-BLAST (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) with the following requirements: optimum primer size of 20 bp, product size between 150 and 500 bp, primer melting temperatures (T<sub>m</sub>) from 57 °C to 63 °C with optimum at 60 °C, maximum 5 bp complementarity between primer pairs and primer location at least 2 bp away from targeted microsatellite locus. Each primer pair (forward and reverse) was tested

**Table 2.** Characteristics of the 17 polymorphic microsatellite loci developed for *Ganoderma boninense*

Locus	GeneBank number	Repeat motif	Primers Sequences (5'-3')	F Start/ R End (bp)	Tm (°C)	Size (bp)
3d	KT124389	(CA) <sub>8</sub>	F:AGCCCGCAGAGCCAATGTCC R:GGCGTGCGACTGCTGGTTCA	7382 7584	59.41 60.25	197-225
6d	KT124390	(AG) <sub>8</sub>	F:ACCTGAGGCGCTGAAAAGCCA R:TGACCAGCCCCACCGTCCTC	9509 9817	58.97 60.18	319-331
7a	KT124391	(CAG) <sub>6</sub>	F:TCGGGTAGGCTCGCAGGTGG R:GGGCCGCACAGGTCGAGAAA	4916 5245	60.32 59.35	338-356
11a	KT124392	(CCGA) <sub>5</sub>	F:GCGTGATGAAGTCGGGCGGG R:GCAGCCGGGATAGTGCGTGT	2636 2883	60.45 59.45	248-284
11d	KT124393	(AGG) <sub>8</sub>	F:TCTGCGGTGGGGGTTGTGCT R:GTCATCGCATCGGGGGTCGG	8190 8456	60.68 59.63	274-289
17b	KT124394	(AGC) <sub>19</sub>	F:CGGGAAGTGGTGAACGGTGGT R:GGGTGGCTTGACAGCGGCAT	7604 7820	58.98 59.97	234-243
18a	KT124395	(CAC) <sub>6</sub>	F:GCGTTGTGCGATGGACGGGG R:TGCGGAGGTGGATGTGGGT	60050 60301	60.11 59.90	263-293
19a	KT124396	(AG) <sub>6</sub>	F:GGCCGAGCCGAAAACGTCACA R:GGGCGGAGCGAGTCAGTTGT	62519 62717	60.24 59.08	208-216
20d	KT124397	(GCA) <sub>9</sub>	F:CGCCATGCCACCACAGAG R:GACCCGGCTGCCCGAATGAG	6268 6545	60.04 59.84	283-325
22a	KT124398	(CAG) <sub>6</sub>	F:TCGGAAGGCAGACGGACGGA R:CTAGGCGTGCGTGGCAGAGC	67367 67565	59.62 60.45	205-217
24d	KT124399	(CAG) <sub>11</sub>	F:GCACAGGCACAAGCGCAAGG R:CGACGACCGCCCCAAAGGAT	10128 10359	59.71 59.14	204-267
33a	KT124400	(CCA) <sub>9</sub>	F:AGCTCCCCTCCAGCTCCAAC R:GAATGCGGCGGGGAAACGGA	43018 43188	59.90 60.04	171-186
37d	KT124401	(CGG) <sub>8</sub>	F:TCCACGCCTGAGGGCCAGAA R:TCGGAGTCGAAGCCACCCTTCA	7092 7420	59.82 59.60	311-350
42d	KT124402	(GT) <sub>9</sub>	F:ACAAGGCTCAAGGCAGCGCA R:GCACACCCCAGCAACAGGAGG	17415 17621	59.55 59.91	212-224
47d	KT124403	(GA) <sub>10</sub>	F:GGCGACGAGGGCACGAGAGA R:CCGCACTTTCGCCAACCAACC	392 655	60.66 59.09	273-297
51d	KT124404	(CGG) <sub>9</sub>	F:CGAACGCCATCCGAGCAACAGA R:TTGGCTCAAAGCCCCGCAAA	8504 8782	59.80 57.64	281-317
52d	KT124405	(GA) <sub>8</sub>	F:CCTTGTGGTGGTAGGCGGCA R:ATCTCCGTGGCACCGCATCCT	3911 4212	58.99 60.31	313-327



with in silico PCR, using NJ3 genome sequence and a local Blastn procedure (Altschul et al. 1990) to select primer pairs that match on a single genome location. Selected primer pairs were built with a M13 (5' CACGACGTTGTAAAACGA3') (Boutin-Ganache et al. 2001) tail added on the 5' side of the forward primer. This M13 sequence is complementary to the M13 fluoro-tagged primer.

The first *in vitro* validation of the designed primer pairs was performed on the 5 Indonesian *G. boninense* isolates maintained as pure cultures (the 5 Indonesian isolate set) to identify primer pairs exhibiting mono locus amplification pattern and polymorphism. Only the primer pairs giving mono locus amplification and polymorphic products were used to genotype the set of 107 *G. boninense* fruiting bodies and two *G. resinaceum* isolates. All the primer characteristics are given in the Table 2.

#### PCR amplification and genotyping

Each microsatellite was amplified in simplex 10 µl PCR reactions containing 5 ng DNA, 100 nM each of fluoro-tagged M13 (FAM, NED, VIC, PET), 100 nM reverse primer, 80 nM of forward primer, 0.2 mM dNTPs, 0.6 mM MgCl<sub>2</sub>, 1x reaction buffer and 0.2 U QIAGEN Taq DNA polymerase, 5% DMSO and BSA at 0.8 µg.µl<sup>-1</sup>. After amplification, 2 µl of simplex PCR product for each dye (2 markers per dye) were pooled in a plate and 10 µl of a mixture of Genescan LIZ 600 size standard and formamide (in the ratio 1:125) was added to each well for genotyping.

For SSR validation using the 5 Indonesian isolate set, genotyping was performed on a Li-Cor 4300 automated DNA analyzer with IRDye size standards with 50 to 350 bp and 50 to 700 bp (Li-Cor Inc.). For SSR genotyping of the 107 *G. boninense* samples, PCR products were analyzed using a 3500xl DNA Analyzer (Applied Biosystems). All alleles were scored using Genemapper (version 4.1, Applied Biosystems).

#### SSR characteristics and diversity analysis

Allele number and genetic diversity are key marker characteristics as they quantify marker information content. For each locus, allele number, observed ( $H_o$ ) and expected heterozygosity ( $H_e$ ) were calculated using Genetix 4.05.2 (Belkhir et al. 2004) for all samples together and for each geographical area: peninsular Malaysia, Borneo and Sumatra. Polymorphism Information Content (PIC) was calculated using the R package PopGenkit that is based on Bolstein et al. (1980). Moreover, number of observed genotypes was calculated for each locus, both for the three geographical areas and for the whole areas.

Linkage disequilibrium (LD) tests for and quantifies the statistical correlation between loci or alleles and, in particular, can detect whether markers convey redundant information due to their physical linkage. LD analysis between allele pairs was conducted at two levels, first on the whole set of samples and second for each geographical area, using Fstat 2.9.3. Significance was assessed using a permutation test accounting for Bonferroni corrections with 40800 permutations (<http://www2.unil.ch/popgen/softwares/fstat.htm>, developed by Goudet (2002)).

A potential drawback of using SSRs is the presence of null alleles that, for various reasons (Dakin et al. 2004), fail to amplify up to detection level. Such null alleles generate false homozygote genotypes and may cause bias in population genetic estimates (Pemberton et al. 1995, Chapuis et al. 2007). For example, in genetic structure analysis, null alleles may decrease accuracy in assigning samples to populations and overestimate  $F_{st}$  (Carlsson 2008). It is, therefore, necessary to identify loci with such null alleles and to discard them for further analysis. Null alleles can be identified using indirect methods that account for excess homozygotes at null allele loci compared to other loci. As recommended by Dabrowski et al (2014) we combined the results of two programs based on indirect methods, ML-NullFreq (Kalinowski et al. 2006) and Microchecker (Van Oosterhout et al. 2004) to determine loci with null alleles. These two programs are complementary in the sense that Microchecker has a high detection rate associated with a high false positive rate while ML-NullFreq is less powerful and detect less false positive. Null allele detection was carried out for the three geographical areas independently and we considered that loci with at least five negative results over six (three populations tested with two softwares) could be considered as not having null alleles and kept for further analysis.

A descriptive analysis of the genetic structure of the samples was carried out using ade4 R package (Dray et al. 2007). To measure possible genetic differentiation between the geographical areas,  $F_{st}$  was calculated between samples from the three geographical areas and a principal component analysis (PCA) performed to study possible sample differentiation. The calculation of  $F_{st}$  was based on Nei's genetic distance (Nei 1973) with heterozygosities weighted by group sizes and significance assessed with a standard permutation test based on Markov chain (1000 permutations). Both PCA and  $F_{st}$  calculations analysis were done twice, first with the SSR set with no null alleles and then with the full SSR set.

## Results

#### Genome sequencing and assembly

454 pyrosequencing of the NJ3 isolate generated 2,599,370 reads with an average read length of 378 bp. Illumina sequencing generated 276,922,790 reads with an average read length of 51 bp. After assembly with the CLC Genomics Workbench 4.6.1 software, 21,425 contigs were obtained with lengths varying from 501 bp to 89,905 bp and a mean of 2,864 bp. This assembly generated a genome total length of 63,034,190 bp. The N25, N50 and N75 lengths were 10,965 bp, 5,913 bp and 2,563 bp respectively.

**Table 3.** Repartition and characteristics of detected microsatellite loci

Motif	Total number of detected markers	Number of repeats	
		min	max
Di	1630	5	26
Tri	731	5	24
Tetra	50	5	16
Penta	18	5	7
Hexa	58	5	14

#### Characterization of microsatellite markers in *G. boninense*

Based on the draft genome assembly, MSATCOMMANDER software identified a total of 2,487 candidate SSR markers in the 21,425 contigs. The number of loci detected for each type of microsatellite loci and the repeat number decrease from di-nucleotide motif to hexa-nucleotide. However, there is an exception for penta-nucleotide loci which is the motif with the lowest number of detected loci (Table 3).

The 1,687 SSR markers with flanking sequences of at least 400 bp, were selected to facilitate primer design. This selection was fairly equitable for most of the types of SSR, with a mean of 32.5% of loci discarded for di-, tri-, tetra- and hexa-nucleotide SSR markers, 61% of the candidate penta-nucleotide SSR markers. To save time, 145 SSR markers were selected randomly for designed primer pairs for further assessment of the marker properties. All the primer pairs were designed to produce an expected PCR product size ranging from 150 bp to 400 bp. After *in silico* PCR, only 110 primer pairs showed the expected mono-locus amplification pattern. After *in vitro* PCR using these 110 primer pairs, 17 primer pairs (15.5%) showed no amplification while 93 primer pairs (84.5%) showed reproducible mono-locus amplification. Among these 93 primer pairs, 46 (49.5%) exhibited polymorphic amplification products from the 5 Indonesian isolates. The SSR loci amplified by these 46 primer pairs were used as markers to genotype the collection of 107 *G. boninense* isolates and the two *G. resinaceum* isolates. After this step, only 17 SSR markers showed proper amplification, i.e. migration patterns characterized by a main peak over 200 RFU (Relative Fluorescent Unit), no more than two amplified peaks, and a maximum of 10% missing data. The primer pairs were designed with a wide range of expected PCR products length, so for each dye, two markers (differing at least 16 nucleotides) can be multiplexed for genotyping (i.e. 8 markers per plate), increasing genotyping speed for large scale analyses. All the properties of the 17 properly amplified polymorphic SSR markers are summarized in Table 2. This set comprises 6 di-nucleotides, 10 tri-nucleotides and one tetra-nucleotide with repeat numbers from 5 to 11 and PCR products ranging from 171 bp to 356 bp. All primers have theoretical melting temperature between 57.6 °C and 60.7 °C and a difference between forward and reverse primers of no more than 2.2 °C (Table.2).

#### SSR characteristics and diversity analysis

We identified 136 alleles for the 17 SSR set when genotyping the 107 *G. boninense* isolates. The allele number per locus varied from 3 (locus 17b) to 14 (locus 24d) with an average value of 7 (Table 4). The number of identified genotypes overall geographic areas varied from 5 (locus 17b) to 38 (locus 20d) with an average of 17 (Table 4). No PCR amplification failure was observed for five loci (locus 7a, 22a, 33a, 42d, 51d) whereas there was up to 8.4% failure for locus 18a. Difference between mean observed ( $H_o$ ) and expected heterozygosity ( $H_e$ ) was not significant for any geographical areas. PIC values for the 17 SSR markers ranged between 0.298 (locus 17b) and 0.813 (locus 20d). LD was significant ( $p$ -value < 0.01) between alleles for four pairs of loci in the full set of samples: 18a and 47d, 19a and 20d, 3d and 20d, 11d and 47d (ranging from 0.001 to 0.009). All these markers belonged to different contigs. For the various geographical areas, LD was only significant for the pair 19a and 20d in the peninsular Malaysia samples (0.002) and the pair 3d and 20d in the Sumatra samples (0.001). Based on our decision rules, 9 loci (7a, 22a, 33a, 6d, 11d, 20d, 37d, 47d and 51d) showed no convincing indications of null alleles (see supplementary information Table 1) whereas 8 loci (11a, 18a, 19a, 17b, 3d, 24d, 42d, 52d) showed evidence for null alleles presence.

#### Genetic structure

**Table 4.** Genetic diversity characteristics of the 17 *Ganoderma boninense* SSR set<sup>a</sup>

Locus	Overall			Indonesia, n = 47					Malaysian Borneo, n = 20					Peninsular Malaysia, n = 40				
	N	N <sub>G</sub>	PIC	N <sub>p</sub>	N <sub>A</sub>	N <sub>GP</sub>	H <sub>O</sub>	H <sub>e</sub>	N <sub>p</sub>	N <sub>A</sub>	N <sub>GP</sub>	H <sub>O</sub>	H <sub>e</sub>	N <sub>p</sub>	N <sub>A</sub>	N <sub>GP</sub>	H <sub>O</sub>	H <sub>e</sub>
07a	7	13	0.408	47	5	8	0.2979	0.4136	20	2	3	0.3500	0.4090	40	6	10	0.3500	0.4541
11a	10	22	0.686	47	5	11	0.5106	0.7156	13	8	9	0.4615	0.8400	39	8	15	0.7179	0.7696
18a	11	27	0.711	47	10	18	0.3191	0.6156	17	6	6	0.0588	0.7968	34	9	16	0.3235	0.7401
19a	5	10	0.545	47	4	8	0.3191	0.6083	20	3	4	0.4500	0.4218	39	5	9	0.5641	0.6810
22a	5	10	0.54	47	5	8	0.5957	0.5749	20	3	4	0.3500	0.3679	40	4	7	0.5750	0.5991
33a	6	13	0.575	47	5	8	0.6170	0.5758	20	5	8	0.8000	0.7590	40	5	9	0.5000	0.6434
17b	3	5	0.298	45	2	3	0.1111	0.4057	20	2	3	0.1500	0.4500	40	3	4	0.1000	0.2291
03d	10	20	0.771	47	7	16	0.5319	0.8065	19	6	11	0.5263	0.7980	40	7	13	0.5250	0.7889
06d	6	14	0.658	47	6	11	0.6596	0.7008	19	4	6	0.4737	0.5277	36	4	8	0.5278	0.6553
11d	6	16	0.568	47	5	10	0.4681	0.5706	20	4	8	0.8000	0.7577	39	6	9	0.4872	0.5311
20d	13	38	0.813	47	13	26	0.8298	0.8296	19	7	12	0.8421	0.8307	40	11	22	0.8000	0.8468
24d	14	29	0.79	46	9	17	0.4130	0.8067	20	8	12	0.4000	0.8346	40	11	18	0.4500	0.7801
37d	9	14	0.505	47	6	9	0.3830	0.4310	20	6	7	0.5500	0.5205	38	5	9	0.5263	0.6281
42d	5	8	0.411	47	4	6	0.1064	0.4482	20	3	3	0.3000	0.3038	40	4	7	0.1750	0.4402
47d	11	21	0.584	47	8	13	0.4894	0.6719	18	6	9	0.6111	0.7206	40	6	8	0.3750	0.5589
51d	8	22	0.717	47	8	15	0.7234	0.7081	20	6	7	0.6500	0.5590	40	7	16	0.7000	0.7924
52d	7	15	0.694	47	6	10	0.3404	0.6495	18	6	9	0.2222	0.8111	39	6	11	0.3077	0.7263
Mean	8	17	0.604	46	6	12	0.4386	0.5986	19	5	7	0.4325	0.5959	39	6	11	0.4360	0.5980

<sup>a</sup> N number of allele per locus; N<sub>G</sub> Number of identified genotypes per locus over all populations; PIC: Polymorphism Information Content based on Bolstein's et al. method; N<sub>p</sub> Number of sample positively amplified per locus per population; N<sub>A</sub> number of allele per locus per population; N<sub>GP</sub> Number of identified genotypes per locus per population; H<sub>O</sub> Observed heterozygosity; H<sub>e</sub> Expected heterozygosity



The main result shown by PCA (first axis, eigenvalue 4.36) is the clear separation between samples from Borneo and those from Sumatra and peninsular Malaysia. Including or excluding loci with null alleles does not affect this result (first axis eigenvalue 4.8% see Figures 1 A and B). Samples collected from peninsular Malaysia and Sumatra are interspersed within a homogenous cluster, indicating that they belong to the same genetic group. The PCA results are in agreement with pairwise  $F_{st}$  values calculated between the samples from the three geographical areas. For the 9 markers without null alleles,  $F_{st}$  was 0.0792 ( $p$ -value 0.0009) between Sumatra and Borneo samples and 0.0773 ( $p$ -value 0.0009) between peninsular Malaysia and Borneo samples but only 0.0111 ( $p$ -value 0.1778) between Sumatra and peninsular Malaysia samples (Table 5). For the 17 marker set, the results were similar (Table 5) with slightly smaller but still significant  $F_{st}$  values of 0.056 between Borneo and Sumatra samples and 0.061 between Borneo and peninsular Malaysia samples.

**Table 5.**  $F_{st}$  calculation based on 3 populations

	Sumatra/Peninsular Malaysia	Sumatra/Borneo	Peninsular Malaysia/Borneo
9 Markers	0.0111	0.0792**	0.0773**
17 Markers	0.0133	0.0564**	0.061**

Pairwise  $F_{st}$  based on Nei (1973)'s method; 9 Markers figures the set containing markers with a negative result to both allele null test at 99%; 17 Markers figures the total set of microsatellite markers; \*\*significant value for  $p$ -value < 0.01 after 1000 permutation.

#### Cross-species transferability of SSR marker set

All the 17 SSR markers could be amplified for the *G. resinaceum* strain CBS19476 and 16 of the SSR markers could be amplified for the *G. resinaceum* strain CBS22036 for which locus 18a did not yield any PCR products. For the whole 17 SSR marker set, only two new alleles were discovered on the *G. resinaceum* isolates. Locus 37d had a specific allele for strain CBS22036 only with a size of 326 bp, which is within the range of detected allele sizes in *G. boninense* (311-350 bp). Locus 52d had a specific allele for both strains of *G. resinaceum*. This allele was 309 bp, which is outside the range of sizes found for *G. boninense* isolates (313-327 bp).

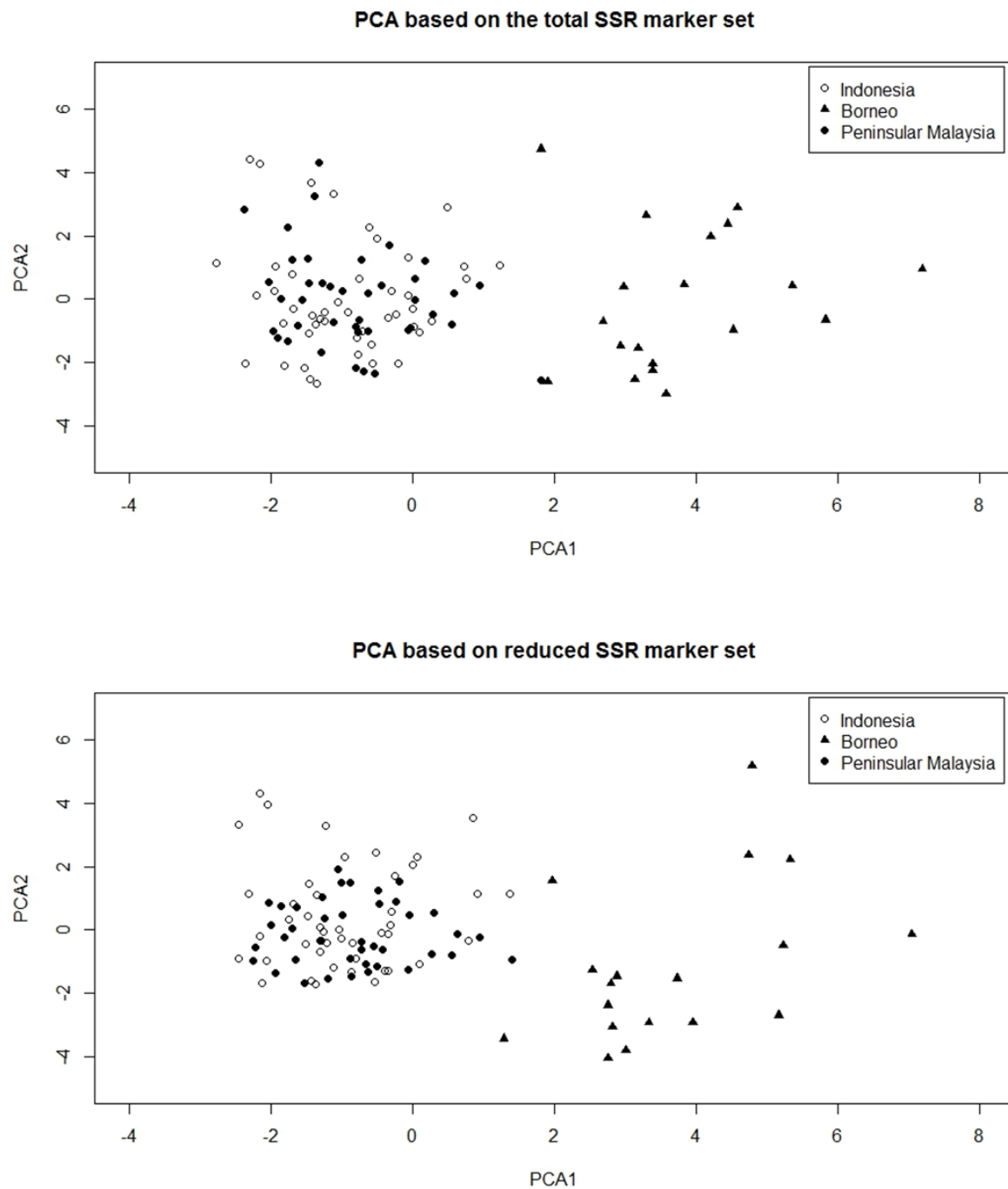
## Discussion

This study identified 2487 SSR loci in a *G. boninense* draft genome assembly and, focusing on the fastest way to obtain a workable minimum microsatellite set for genetic diversity analysis, 17 SSR markers were developed. These markers exhibited satisfactory amplification and polymorphism properties for large scale population genotyping. To reduce the time and cost of genotyping, these SSR markers were designed to be used in multiplex. This is the first set of locus specific SSR markers developed for *G. boninense*. Until now, only random amplified microsatellites (RAMS) (Zakaria et al. 2005), rapid amplification of polymorphic DNA (RAPD) (Abu-Seman et al. 1996, Zakaria et al. 2005) and internal transcribed spacer (ITS) (Moncalvo et al. 1995 b, Smith et al. 2000, Utomo et al. 2005) markers have been developed for the pathogenic strains of *Ganoderma* species found on oil palms.

Our study confirms the advantages of using assembled reads (Seq-Assembly-SSR) rather than raw reads for the detection and development of SSR markers for non-model species, as already shown by Cai et al. (2013). We were able to design primers for 100% of the selected SSRs, compared to 49.6% for studies based on 454 raw reads and 30 to 40% for studies based on Illumina raw reads. Among the 145 primer pairs tested *in vitro*, 93 (84.5%) showed mono locus amplification, which is comparable to Cai et al. (2013) who obtained 90.7% locus specific amplification from primer pairs designed using a genome assembly. The effectiveness of the Seq-Assembly-SSR approach in targeting locus specific markers and primer design is the result of two improvements. Firstly, assembled contigs provide longer sequences defining primers with an average contig length of 2,864 bp after assembly compared to raw reads of 51 bp for the Illumina used or 378 bp for the 454. Secondly, sequence redundancy is reduced by the assembly process: while the amount of raw reads was around 4 Gb, the assembly process reduced the total amount of data to around 61 Mb. This reduction in redundancy improves the efficiency of pre-screening for primers amplifying single locus PCR products using *in silico* PCR (Cai et al. 2013). Furthermore, the selection of microsatellite loci with flanking sequences of at least 400 bp facilitates primer design and multiplexed genotyping which reduces genotyping costs for large sets of samples.

The number of SSRs detected in the *G. boninense* genome draft assembly is within the range of numbers of SSRs detected in fungal genomes as reviewed by Karaoglu et al. (2005) and which ranges from 124 (*Encephalitozoon cuniculi*) to 8376 (*Neurospora crassa*). However, the number of SSRs detected in this study is

**Figure 1. Global PCA results based on two set of markers.** (A) Global genetic diversity observed by PCA analyses based on the total SSR marker set (B) Global genetic diversity observed by PCA analyses based on the reduced 9 SSR marker set including only markers without null allele predicted.



almost twice the number of SSRs identified in *G. lucidum* (1584) by Qian et al. (2013). As our *G. boninense* genome assembly is only a draft, no firm conclusions can be drawn. However, a comparison of the estimated genome size of these two fungi (43 Mb for *G. lucidum* and 65 Mb estimated for the *G. boninense* draft assembly), as well as the much larger number of SSRs, suggests that there may have been a large duplication or deletion event at genome scale in the *Ganoderma* genus.

While most studies recommend a minimum of 8 to 11 SSR markers to perform sound statistical analyses for population genetics (van Asch et al. 2010), this study developed 17 SSR markers with good amplification and polymorphism properties that should allow comprehensive large scale population studies in *G. boninense*. Fungus population studies have used on average 10 SSR markers and generally no more than 16 SSR markers (Schoebel et al. 2013, Wang et al. 2014). Of the 17 SSR markers developed in this study, mean PIC value is about 0.604 and among them 14 presented a PIC value greater than 0.5 indicating that they are highly informative (PIC >0.5) according to Botstein et al. (1980). Compared to the mean PIC value of 0.33 obtained for SSRs described on another *Polyporales* by Zhang et al. (2015), our SSRs show high polymorphic information. Among our 17 SSRs, 9 had no null alleles according to criteria used in this study. While the LD detected at the whole sample scale could result from sample structure, LD was significant for only two pairs of alleles within one population out of three, indicating that the information conveyed by each SSR marker is non redundant. The visualization of putative relationships among investigated samples by PCA showed clear differentiation of a group of samples from Borneo apart of a mixed group of samples from Sumatra and peninsular Malaysia. PCA results were confirmed by  $F_{st}$  values: about 0.07 between Borneo and Sumatra or Borneo and peninsular Malaysia, while only about 0.01 between Sumatra and peninsular Malaysia. PCA and  $F_{st}$  results are similar whether the full 17 SSR set or the set of 9 SSR with no null alleles is used. This indicates a potential high gene flow between the Sumatra and peninsular Malaysia populations, even though they are separated by the Malacca Strait, and that the Borneo population is partially isolated.

The amplification of 16 of the SSRs for 2 *G. resinaceum* isolates, a species closely related to *G. boninense* (Moncalvo et al. 1995 a, Moncalvo et al. 1995 b, Binder et al. 2013) raises the questions of species discrimination using this SSR set and species delimitation within the *Ganoderma* genus. It is often considered that SSRs are poorly transferable between fungal species (Dutech et al. 2007, Schoebel et al. 2013) except in some particular cases involving species complexes (Perez et al. 2012). The non-transferability of SSR loci between *G. lucidum* and *G. boninense* could be however verified by comparing the flanking sequences of the SSRs we detected for *G. boninense* to sequences in the *G. lucidum* genome (Chen et al. 2012), but it was impossible to identify motifs of at least 20 bp that could be to design trans-specific SSR primers. Finally, the results for *G. resinaceum* are too limited to draw any firm conclusion: more individuals from *G. resinaceum* and the sequencing of fungi barcode markers such as ITSs or tubulin genes (Moncalvo et al. 1995 a, Moncalvo et al. 1995 b) would be necessary to determine the species status of these two taxa.

**Acknowledgments** We should like to thank all the PT Socfin Indonesia team for the sampling in Indonesian plantations and especially Mr Indra, Mr Sri, Mr Turnbull and Mr Williams. We should also like to thank all the FELDA team for the sampling in the Malaysian plantations.

## References

- Abdelkrim, J., B. Robertson, J. A. Stanton and N. Gemmell (2009). Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *Biotechniques* 46(3): 185–92.
- Abu-Seman, I., M. Thangavelu and T. Swinburne (1996). The use of RAPD for identification of species and detection of genetic variation in *Ganoderma* isolates from oil palm, rubber and other hardwood trees. Proceeding of Porim International palm oil congress, Kuala Lumpur, Malaysia.
- Ali, S., P. Gladieux, H. Rahman, M. S. Saqib, M. Fiaz, H. Ahmad, M. Leconte, A. Gautier, A. F. Justesen, M. S. Hovmoller, J. Enjalbert and C. de Vallavieille-Pope (2014). Inferring the contribution of sexual reproduction, migration and off-season survival to the temporal maintenance of microbial populations: a case study on the wheat fungal pathogen *Puccinia striiformis* f.sp. *tritici*. *Mol Ecol* 23(3): 603–17.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). Basic local alignment search tool. *J Mol Biol* 215(3): 403–10.
- Barbara, T., C. Palma-Silva, G. M. Paggi, F. Bered, M. F. Fay and C. Lexer (2007). Cross-species transfer of nuclear microsatellite markers: potential and limitations. *Mol Ecol* 16(18): 3759–67.
- Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste and F. Bonhomme (2004). genetix. Populations Genet Software Windows TM Univ Montp 4(05).

- Binder, M., A. Justo, R. Riley, A. Salamov, F. Lopez-Giraldez, E. Sjökvist, A. Copeland, B. Foster, H. Sun, E. Larsson, K. H. Larsson, J. Townsend, I. V. Grigoriev and D. S. Hibbett (2013). Phylogenetic and phylogenomic overview of the *Polyporales*. *Mycologia* 105(6): 1350–1373.
- Botstein, D., R. L. White, M. Skolnick and R. W. Davis (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32(3): 314–31.
- Boutin-Ganache, I., M. Raposo, M. Raymond and C. F. Deschepper (2001). M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allelizing methods. *Biotechniques* 31(1): 24–+.
- Breton, F., Y. Hasan, S. Hariadi, Z. Lubis and H. De Franqueville (2006). Characterization of parameters for the development of an early screening test for basal stem rot tolerance in oil palm progenies. *Journal of Oil Palm Research*: 24–36.
- Cai, G., C. W. Leadbetter, M. F. Muehlbauer, T. J. Molnar and B. I. Hillman (2013). Genome-wide microsatellite identification in the fungus *Anisogramma anomala* using Illumina sequencing and genome assembly. *PLoS One* 8(11): e82408.
- Carlsson, J. (2008). Effects of Microsatellite Null Alleles on Assignment Testing. *Journal of Heredity* 99(6): 616–623.
- Castoe, T. A., A. W. Poole, W. Gu, A. P. Jason de Koning, J. M. Daza, E. N. Smith and D. D. Pollock (2010). Rapid identification of thousands of copperhead snake (*Agkistrodon contortrix*) microsatellite loci from modest amounts of 454 shotgun genome sequence. *Mol Ecol Resour* 10(2): 341–7.
- Chapuis, M. P. and A. Estoup (2007). Microsatellite null alleles and estimation of population differentiation. *Mol Biol Evol* 24(3): 621–31.
- Chen, S., J. Xu, C. Liu, Y. Zhu, D. R. Nelson, S. Zhou, C. Li, L. Wang, X. Guo, Y. Sun, H. Luo, Y. Li, J. Song, B. Henrissat, A. Levasseur, J. Qian, J. Li, X. Luo, L. Shi, L. He, L. Xiang, X. Xu, Y. Niu, Q. Li, M. V. Han, H. Yan, J. Zhang, H. Chen, A. Lv, Z. Wang, M. Liu, D. C. Schwartz and C. Sun (2012). Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. *Nat Commun* 3: 913.
- Corley, R. H. V. and P. B. Tinker (2003). *The oil palm*. Oxford ; Malden, MA, Blackwell Science.
- Dabrowski, M. J., S. Bornelov, M. Kruczyk, N. Baltzer and J. Komorowski (2014). 'True' null allele detection in microsatellite loci: a comparison of methods, assessment of difficulties and survey of possible improvements. *Mol Ecol Resour* 15(3): 477–88.
- Dakin, E. E. and J. C. Avise (2004). Microsatellite null alleles in parentage analysis. *Heredity (Edinb)* 93(5): 504–9.
- Doyle, J. J. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem bull* 19: 11–15.
- Dray, S. and A.-B. Dufour (2007). The ade4 package: implementing the duality diagram for ecologists. *Journal of statistical software* 22(4): 1–20.
- Durand-Gasselin, T., H. Asmady, A. Flori, J. C. Jacquemard, Z. Hayun, F. Breton and H. de Franqueville (2005). Possible sources of genetic resistance in oil palm (*Elaeis guineensis* Jacq.) to basal stem rot caused by *Ganoderma boninense*--prospects for future breeding. *Mycopathologia* 159(1): 93–100.
- Dutech, C., B. Barres, J. Bridier, C. Robin, M. G. Milgroom and V. Ravigne (2012). The chestnut blight fungus world tour: successive introduction events from diverse origins in an invasive plant fungal pathogen. *Mol Ecol* 21(16): 3931–46.
- Dutech, C., J. Enjalbert, E. Fournier, F. Delmotte, B. Barres, J. Carlier, D. Tharreau and T. Giraud (2007). Challenges of microsatellite isolation in fungi. *Fungal Genet Biol* 44(10): 933–49.
- Faircloth, B. C. (2008). msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Mol Ecol Resour* 8(1): 92–4.
- Flood, J., P. D. Bridge and M. Holderness (2000). *Ganoderma* diseases of perennial crops. Wallingford, Oxon England ; New York, CABI.
- Gardner, M. G., A. J. Fitch, T. Bertozzi and A. J. Lowe (2011). Rise of the machines--recommendations for ecologists when using next generation sequencing for microsatellite development. *Mol Ecol Resour* 11(6): 1093–101.
- Goudet, J. (2002). Fstat. A program to estimate and test gene diversities and fixation indices, release 2(1).



- Gregory, T. R., J. A. Nicol, H. Tamm, B. Kullman, K. Kullman, I. J. Leitch, B. G. Murray, D. F. Kapraun, J. Greilhuber and M. D. Bennett (2007). Eukaryotic genome size databases. *Nucleic Acids Res* 35(Database issue): D332–8.
- Ho, Y. and A. Nawawi (1986). Isolation, Growth and Sporophore Development of *Ganoderma boninense* from Oil Palm in Malaysia. *Pertanika* 9(1): 69–73.
- How, Y. and A. Nawawi (1985). *Ganoderma boninense* Pat. from basal stem rot of oil palm (*Elaeis guineensis*) in Peninsular Malaysia. *Pertanika* 8(3): 425–428.
- Idris, A., A. Kushairi, S. Ismail and D. Ariffin (2004). Selection for partial resistance in oil palm progenies to *Ganoderma* basal stem rot. *J Oil Palm Res* 16(2): 12–18.
- Jarne, P. and P. J. L. Lagoda (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution* 11(10): 424–429.
- Kalinowski, S. T. and M. L. Taper (2006). Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. *Conservation Genetics* 7(6): 991–995.
- Karaoglu, H., C. M. Lee and W. Meyer (2005). Survey of simple sequence repeats in completed fungal genomes. *Mol Biol Evol* 22(3): 639–49.
- Miller, R. N. G., M. Holderness, P. D. Bridge, G. F. Chung and M. H. Zakaria (1999). Genetic diversity of *Ganoderma* in oil palm plantings. *Plant Pathology* 48(5): 595–603.
- Moncalvo, J. M., H. F. Wang and R. S. Hseu (1995 a). Gene phylogeny of the *Ganoderma lucidum* complex based on ribosomal DNA sequences. Comparison with traditional taxonomic characters. *Mycological Research* 99: 1489–1499.
- Moncalvo, J. M., H. H. Wang and R. S. Hseu (1995 b). Phylogenetic-Relationships in *Ganoderma* Inferred from the Internal Transcribed Spacers and 25s Ribosomal DNA-Sequences. *Mycologia* 87(2): 223–238.
- Montarry, J., D. Andrivon, I. Glais, R. Corbiere, G. Mialdea and F. Delmotte (2010). Microsatellite markers reveal two admixed genetic groups and an ongoing displacement within the French population of the invasive plant pathogen *Phytophthora infestans*. *Mol Ecol* 19(9): 1965–77.
- Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci U S A* 70(12): 3321–3.
- Pemberton, J. M., J. Slate, D. R. Bancroft and J. A. Barrett (1995). Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Mol Ecol* 4(2): 249–52.
- Perez, G., B. Slippers, M. J. Wingfield, B. D. Wingfield, A. J. Carnegie and T. I. Burgess (2012). Cryptic species, native populations and biological invasions by a eucalypt forest pathogen. *Mol Ecol* 21(18): 4452–71.
- Pilotti, C. A. (2005). Stem rots of oil palm caused by *Ganoderma boninense*: pathogen biology and epidemiology. *Mycopathologia* 159(1): 129–37.
- Pilotti, C. A., F. R. Sanderson and E. A. B. Aitken (2002). Sexuality and interactions of monokaryotic and dikaryotic mycelia of *Ganoderma boninense*. *Mycological Research* 106: 1315–1322.
- Pilotti, C. A., F. R. Sanderson and E. A. B. Aitken (2003). Genetic structure of a population of *Ganoderma boninense* on oil palm. *Plant Pathology* 52(4): 455–463.
- Presti, F. T. and A. P. Wasko (2014). A Review of Microsatellite Markers and their Application on Genetic Diversity Studies in Parrots. *Open Journal of Genetics* 2014.
- Qian, J., H. Xu, J. Song, J. Xu, Y. Zhu and S. Chen (2013). Genome-wide analysis of simple sequence repeats in the model medicinal mushroom *Ganoderma lucidum*. *Gene* 512(2): 331–6.
- Rees, R. W., J. Flood, Y. Hasan and R. M. Cooper (2007). Effects of inoculum potential, shading and soil temperature on root infection of oil palm seedlings by the basal stem rot pathogen *Ganoderma boninense*. *Plant Pathology* 56(5): 862–870.
- Schoebel, C. N., S. Brodbeck, D. Buehler, C. Cornejo, J. Gajurel, H. Hartikainen, D. Keller, M. Leys, S. Ricanova, G. Segelbacher, S. Werth and D. Csencsics (2013). Lessons learned from microsatellite development for nonmodel organisms using 454 pyrosequencing. *J Evol Biol* 26(3): 600–11.
- Schoebel, C. N., E. Jung and S. Prospero (2013). Development of new polymorphic microsatellite markers for three closely related plant-pathogenic *Phytophthora* species using 454-pyrosequencing and their potential applications. *Phytopathology* 103(10): 1020–7.

- Smith, B. J. and K. Sivasithamparam (2000). Internal transcribed spacer ribosomal DNA sequence of five species of *Ganoderma* from Australia. *Mycological Research* 104: 943–951.
- Susanto, A., P. S. Sudharto and R. Y. Purba (2005). Enhancing biological control of basal stem rot disease (*Ganoderma boninense*) in oil palm plantations. *Mycopathologia* 159(1): 153–7.
- Travadon, R., K. Baumgartner, P. E. Rolshausen, W. D. Gubler, M. R. Sosnowski, P. Lecomte, F. Halleen and J. P. Peros (2012 a). Genetic structure of the fungal grapevine pathogen *Eutypa lata* from four continents. *Plant Pathology* 61(1): 85–95.
- Travadon, R., M. E. Smith, P. Fujiyoshi, G. W. Douhan, D. M. Rizzo and K. Baumgartner (2012 b). Inferring dispersal patterns of the generalist root fungus *Armillaria mellea*. *New Phytologist* 193(4): 959–969.
- Turner, P. D. (1981). Oil palm diseases and disorders. *Oil palm diseases and disorders*.
- Utomo, C., S. Werner, F. Niepold and H. B. Deising (2005). Identification of *Ganoderma*, the causal agent of basal stem rot disease in oil palm using a molecular method. *Mycopathologia* 159(1): 159–70.
- van Asch, B., R. Pinheiro, R. Pereira, C. Alves, V. Pereira, F. Pereira, L. Gusmão and A. Amorim (2010). A framework for the development of STR genotyping in domestic animal species: Characterization and population study of 12 canine X - chromosome loci. *Electrophoresis* 31(2): 303–308.
- Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills and P. Shipley (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4(3): 535–538.
- Wang, M., F. Xue, P. Yang, X. Y. Duan, Y. L. Zhou, C. Y. Shen, G. Z. Zhang and B. T. Wang (2014). Development of SSR Markers for a Phytopathogenic Fungus, *Blumeria graminis f.sp tritici*, Using a FIASCO Protocol. *Journal of Integrative Agriculture* 13(1): 100–104.
- Wei, N., J. B. Bemmels and C. W. Dick (2014). The effects of read length, quality and quantity on microsatellite discovery and primer development: from Illumina to PacBio. *Mol Ecol Resour* 14(5): 953–65.
- Zakaria, L., H. Kulaveraasingham, T. S. Guan, F. Abdullah and H. Y. Wan (2005). Random amplified polymorphic DNA (RAPD) and random amplified microsatellite (RAMS) of *Ganoderma* from infected oil palm and coconut stumps in Malaysia. *Asia Pac J Mol Biol Biotechnol* 13: 23–34.
- Zalapa, J. E., H. Cuevas, H. Zhu, S. Steffan, D. Senalik, E. Zeldin, B. McCown, R. Harbut and P. Simon (2012). Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. *Am J Bot* 99(2): 193–208.
- Zane, L., L. Bargelloni and T. Patarnello (2002). Strategies for microsatellite isolation: a review. *Mol Ecol* 11(1): 1–16.
- Zhang, Y., Y. Chen, R. Wang, A. Zeng, M. K. Deyholos and J. Shu (2015). Development of microsatellite markers derived from expressed sequence tags of *Polyporales* for genetic diversity analysis of endangered *Polyporus umbellatus*.

**CHAPITRE III :**

**ETUDE DE LA DIVERSITE  
GENETIQUE DE *GANODERMA*  
*BONINENSE***

Au cours du chapitre précédent, nous avons décrit le développement des marqueurs microsatellites nécessaires à l'étude de la diversité génétique de *G. boninense*. Ces marqueurs nous ont permis au cours du premier chapitre de mettre en évidence deux groupes différenciés entre les individus provenant des plantations de Bornéo et ceux provenant de Malaisie péninsulaire et de Sumatra (groupe noté par la suite Malaisie Péninsulaire/Sumatra ou MP/S). Dans ce chapitre nous porterons notre intérêt dans un premier temps sur la diversité et la structure génétique fine au sein du groupe MP/S, puis dans un deuxième temps nous statuerons de manière plus globale sur la diversité et la structure des populations de *G. boninense* au regard de l'ensemble des échantillons acquis au cours de cette thèse.

La zone géographique MP/S est la zone « historique » de développement et d'expansion de la culture industrielle du palmier à huile en Asie du sud-est (Jacquemard 2011). Depuis le début de l'exploitation industrielle du palmier à huile plusieurs générations de palmiers se sont succédé, c'est-à-dire plusieurs cycles complets de culture (du transfert du jeune palmier en plantation jusqu'à son abattage). On parle ici de **génération de plantations** car en règle générale les palmiers au sein d'un même bloc sont plantés en même temps et abattus au même moment. Cette zone est également la région où l'incidence de la BSR est historiquement la plus importante et donc, dans ce contexte, donc une région clef pour la diversité du pathogène *G. boninense* et son adaptation potentielle à son hôte, le palmier à huile. En effet, plus le contact entre un agent pathogène et son hôte est ancien, plus les forces évolutives ont eu le temps de marquer l'évolution du génome de l'agent pathogène et la répartition des individus au sein d'une zone géographique. De plus cette zone faisant l'objet de pratiques culturelles visant à limiter les pertes causées par l'agent pathogène sur les cultures, les pressions de sélection agissant sur de *G. boninense* y sont accrues.

Nous avons tenté au sein de la première partie de ce troisième chapitre de clarifier les modes de reproduction de *G. boninense*, ainsi que l'échelle de ses capacités de dispersion. Nous avons ainsi voulu établir si une structure à plusieurs échelles géographiques (région, plantation) et si des facteurs tels que les antécédents de la plantation, le fonds génétique des arbres ou le nombre de générations de plantations ayant précédé la génération actuelle influençaient l'organisation spatiale des génotypes. Nous avons également essayé de déterminer si l'éventuelle structure spatiale résultait de l'existence de barrières géographiques (par exemple le détroit de Malacca), se traduisant par des groupes génétiques distincts, ou si au contraire des phénomènes d'isolation par la distance étaient à l'œuvre. Pour finir, nous avons testé différents scénarii démographiques afin de déterminer si la rencontre récente de *G. boninense* avec le palmier à huile et ses conséquences démographiques sur *G. boninense* avait impacté le polymorphisme génétique du pathogène. Cette dernière étape devait également permettre d'établir, sur la base du scénario démographique le plus probable, si un saut d'hôte a été réalisé (signature de goulot d'étranglement puis d'expansion démographique) ou si, au contraire, l'introduction du palmier à huile en Asie du sud-est a simplement fourni à *G. boninense* un hôte supplémentaire à coloniser sans nécessité d'adaptation (uniquement signature d'expansion démographique).

Dans la deuxième partie de ce chapitre, nous avons voulu observer à de plus larges échelles la structuration génétique des individus. Dans un premier temps nous avons intégré

au jeu de données précédant l'intégralité des échantillons du sud-est asiatique de notre collection, puis dans un second temps nous y avons ajouté les individus provenant des plantations Africaines.

La première partie de ce chapitre se présente sous forme d'un article qui sera soumis entre la remise du mémoire et la soutenance de la thèse.

## **1. Diversité et histoire démographique des populations de *Ganoderma boninense* au sein des plantations de Sumatra et de Malaisie Péninsulaire**

### **A shared expansion history between Sumatra and peninsular Malaysia for *Ganoderma boninense* in oil palm plantations**

Mercière M., Boulord R., Carasco-Lacombe C., Breton F., Camus-Kulandaivelu L.

#### **Introduction**

Nowadays, fungi are responsible for half of cultivated plant disease (Lepoivre 2003) and can threaten food security. Basidiomycetes are one of the main classes of phytopathogenic fungi and many of them belong to wood rot class, like the famous *Armillaria mellea* responsible of white rot on conifers (Prospero et al. 2008, Baumgartner et al. 2010). White rot disease, are opposed to brown rot by their capacity of lignin degradation, but are also able to degrade cellulose like brown rot. White rot disease incorporate lots of Basidiomycetes including *Ganoderma* genus.

*Ganoderma* genus is represented by almost 250 species distributed all over the world except in the polar areas (Ryvarden 1985). Species of the *Ganoderma* genus are important wood-decaying fungi, occurring on conifers and hardwoods. They are white-rot telluric basidiomycete with the ability to decay both lignin and cellulose and cause many root and stem rots resulting in worldwide crop and tree losses (Martinez et al. 1995, Miller 1995). However, most of *Ganoderma* species are not well characterized and little information on their biology is available. Over the entire genus, only two *Ganoderma* species have raised some interest in the literature: *G. lucidum*, the model medicinal mushroom, and *G. boninense*, the main agent of Basal Stem Rot disease, the most devastating disease of oil palm (*Eleais guineensis*) in South-East Asia. In the contrary of *G. lucidum*, *G. boninense* native area and range of dispersal are still unclear.

*G. boninense* is mainly found in Asia and Oceania on *palmaceae* species (Steyaert 1967), although its exact geographical distribution and host range are unknown. The BSR disease, caused by *G. boninense*, can impact up to 50% of oil palms in a plantation and is responsible for reducing drastically yield and for killing the palms (Corley et al. 2003). Some occurrences of *G. boninense* were also observed on the coconut tree (*Cocos nucifera*) where it is also responsible for butt and root rot (Frison et al. 1993). Moreover, coconut tree was cited several times as one of the possible source of *G. boninense* inoculum in South-East Asia (Turner 1965 a, How et al. 1985, Miller et al. 1999). This disease was first described in 1931 in Malaysian oil palm plantations by Thompson (Ariffin et al. 2000) and is now found in all oil palm plantation areas in South East Asia. In that context, it was described as "the most annihilating disease of fields palms causing significant losses in Southeast Asia" by Turner in 1981 (Turner 1981). Since this day a lot of efforts have been done to study mechanisms of oil palm defense to *G. boninense* infection and to develop resistant or tolerant crop to BSR in order to fight against this threat (Idris et al. 2004, Durand-Gasselin et al. 2005, Tee et al.

2013, Ho et al. 2014). While some studies attempted to better understand *G. boninense* biology and diversity through detailed morphological descriptions (How et al. 1985) as well as sexual reproduction modalities characterization (Pilotti et al. 2002) and genetic diversity quantification with molecular markers (Miller et al. 1999) or with mating type system (Pilotti et al. 2003), genetic diversity and population structure of *G. boninense* across oil palm plantations in South East Asia are still mostly unknown.

In this study, we choose to focus on peninsular Malaysia and Sumatra plantation as this region is the historical area of both industrial oil palm culture and BSR disease expansion in Asia and is very likely a key region for understanding *G. boninense* diversity. As possible host jump had happen, studying impact on genome and population structuration of an adaptation phenomenon on the pathogen must use a place where host and pathogen are in contact for a period longer than human time scales (McDonald et al. 2002). Indeed, the longer the contact between host and pathogen population, the higher chance we have to detect trace of evolution history on the pathogen genome and population structuration. The investigation of genetic diversity at different geographic scale and demographic history of *G. boninense* through PCA, population genetic index and ABC approach, could gave us clues about the influence of oil palm tree plantation on *G. boninense* populations structuration and evolution ( spreading mode, dispersal capacity, population size,...).

Since early XXIst century, plant pathologists realized that a good knowledge about genetic structure of plant pathogen populations enable the development of efficient control strategies (McDonald 1997). The genetic structure of pathogen population can also affect the durability of genetic resistance or can be used to guide breeding programs to lead to durable resistance (McDonald et al. 2002).

Population genetics through the use of molecular marker will bring effective tool to understand factors which could explain current population structuration such as population size, genetic diversity, gene flow or migration (Prospero et al. 2008) and spreading system (Rees et al. 2012). As pathogens evolve endlessly, adapting themselves to changes of their environment in order to survive, it is important to understand actual pathogen population structure and how those factors influenced this structuration to predict potential evolution of the pathogen populations. In order to sharpen prediction of pathogen population evolution, it is relevant to explore past evolution history of the pathogen. The study of population demographic history throw actual genetic diversity was made possible through the development algorithms based on approximate Bayesian computation (ABC). This could give us some clues about past demographic history of plant pathogen and determined factors that influenced the most demographic evolution history of plant pathogen.

This study is intending to increase our knowledge on *G. boninense* populations in a part of South East Asia oil palm plantations. In the aim of studying those populations we used microsatellite markers developed in a previous study (Mercière et al., 2015) to genotypes fruiting bodies collected on oil palms from several plantations spread over peninsular Malaysia and Sumatra. In this study we intend to answer to the following questions: (i) Is there a difference between individuals from close plantations? (ii) Is there gene flow between peninsular Malaysia and Sumatra despite of the Malacca strait? (iii) Is the adaptation of *G. boninense* to a new host has let some traces on actual *G. boninense* genetic diversity? Answers these questions should help to us to understand disease dynamic and plantation managers to adapt their disease management.

## Material and Methods

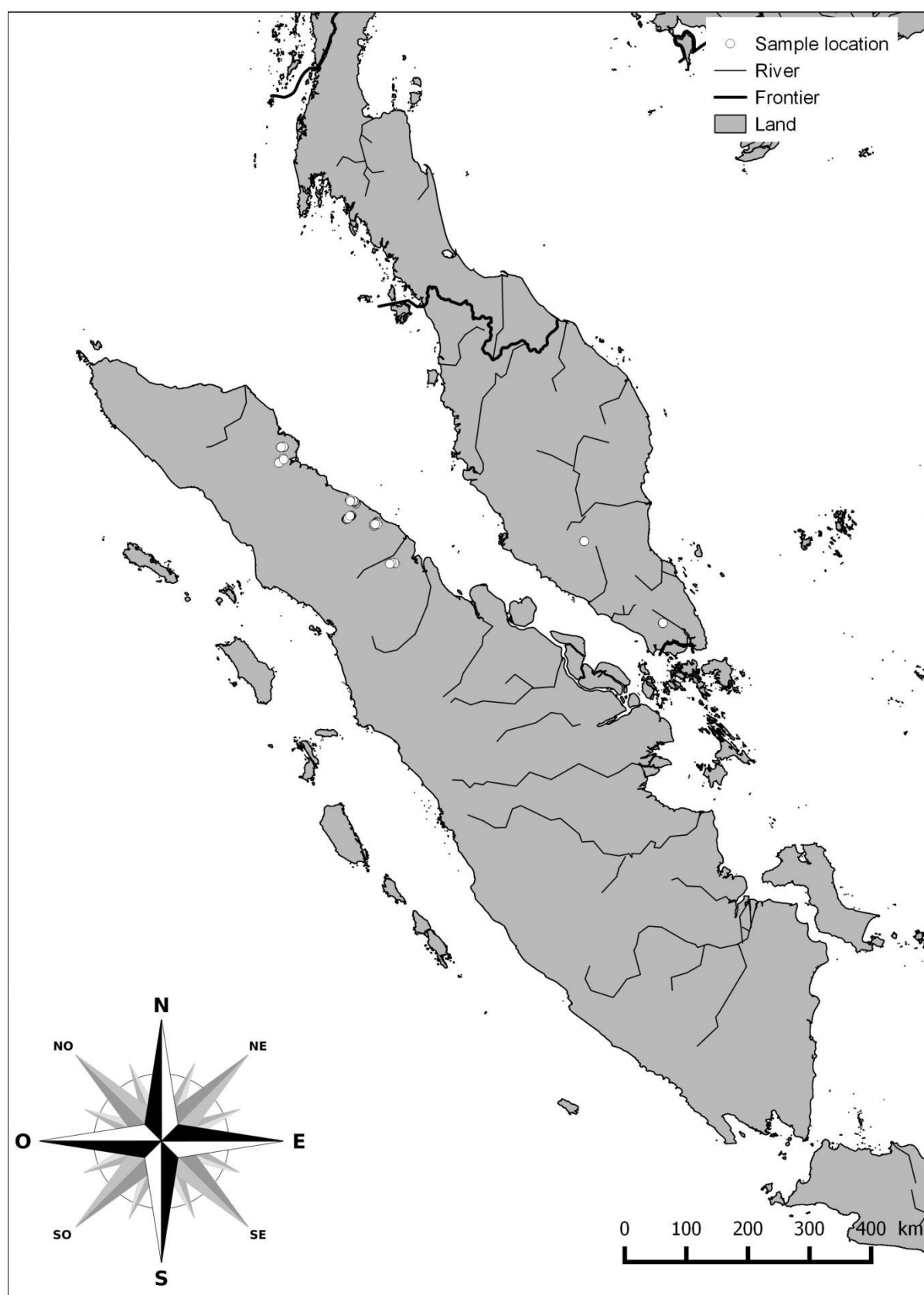


Figure 1. Map presenting Sumatra and peninsular Malaysia sampling



## Samples dispersal and sampling scheme

During year 2013, we sampled 357 *Ganoderma boninense* fruiting bodies growing on oil palm from two FELDA GLOBAL VENTURE plantations in Peninsular Malaysia (Johor and Negeri Sembilan) and from five PT SOCFINDO plantations in Sumatra (Bangun Bandar, Mata Pau, Tanah Gambus, Sei Liput and Aek Loba). A single carpophore was collected on each infected oil palm and was considered as one isolate. Each plantation was explored taking account disease incidence for each block to increase the probability of observing disease symptoms and fruiting bodies. Moreover, sampled blocs were selected to cover plantations as well as possible (North to South, East to West) and, within each bloc, fruiting bodies were collected randomly over the whole block surface. We collected 92 fruiting bodies in Peninsular Malaysia: 45 in Johor, distributed among 3 blocks and 47 in Negeri Sembilan, distributed among 2 blocks. We collected 265 fruiting bodies in Sumatera: 107 in Bangun Bandar (13 blocks), 62 in Mata Pau (9 blocks), 61 in Tanah Gambus (10 blocks), 27 in Sei Liput (12 blocks) and 8 in Aek Loba (5 blocks). For all fruiting bodies, we recorded GPS coordinates (deg°min'sec''N, deg°min'sec''W), number of planting generation and, when possible, planting year and genetic tree background (supplementary material, Table 1). Global repartition of samples is presented in Figure 1. Collected fruiting bodies were surface cleaned with 70 % ethanol and cut to keep only inner part to limit contamination. Each fruiting body piece was freeze-dried and then stored in silica gel. Each fruiting body was considered as one isolate (or sample) and species affiliations of each isolate was based on morphological traits as described in Ho et al. (1986) and lately in Flood et al. (2000 a).

## DNA extraction, PCR amplification and genotyping

We extracted DNA from a total of 357 freeze dried fruiting bodies using Mercière et al. (2015)'s protocol. The 17 microsatellites markers described by Mercière et al. (2015) were used to genotype our 357 samples. Microsatellites markers characteristics are resumed in Table 1. Each microsatellite was amplified in 384 well plates with a volume of 10 µl per PCR reactions following composition described by Mercière et al. (2015). The PCR program was developed following a classical Touch-Down protocol: first step of Taq activation 4 min at 94 °C, 25 cycles constituted of 45 s at 94 °C, 60 °C to 55 °C (with a decrease of 0.5 °C each cycle during 10 cycles) and 1 min 15 s at 72°C. To finalize elongation a last step of 30 min at 72°C was performed. PCR products were analyzed on a 3500xl DNA Analyzer (Applied Biosystems). Then, all alleles were scored using Genemapper (version 4.1, Applied Biosystems). Only missing data free samples were considered for further analysis.

## Genetic analysis and population structuration of *Ganoderma boninense*

We used GenClone 2.0 (Arnaud-Haond et al. 2007) to identify duplicated genotypes. Such duplicated genotypes, that exhibit same allelic combinations for all 17 locus, may correspond either to a same large mycelium sampled on several palm trees or to distinct mycelium resulting from clonal multiplication. As clones are most of the time not accounted for in most of evolution theory and models (Arnaud-Haond et al. 2007) such duplicated genotypes may impact population genetic analysis. Therefore, redundant genotypes were removed.

As SSR loci with null alleles may also cause bias in population genetic analysis (Paetkau et al. 1995, Chapuis et al. 2007), we searched for them using ML-NullFreq (Kalinowski et al. 2006) and Microchecker (Van Oosterhout et al. 2004) softwares, as recommended by Dabrowski et al. (2014) and applied the decision rule described by Mercière et al. (2015). SSRs exhibiting evidence for null allele were removed from population genetics analysis.

Mean allele number, as well as observed (Ho) and expected heterozygosity (He) were calculated at different scale: firstly by plantation and secondly by region using Genetix 4.05.2

(Belkhir et al. 2004). Hardy-Weinberg equilibrium (HWE) was assessed for each locus individually and for all loci globally using Adegnet R package (Jombart 2008) with a classical permutation test for which we performed 1000 permutations. Allelic richness (A) was calculated at different level : Block, plantation and region using PopGenReport R package developed by Adamack et al. (2014) which minimize influence of sampling size by using El Mousadik et al. (1996) method's. We grouped isolates by block and explore potential difference between A values obtained for blocks grouped by planting generation, by plantation, and by region with a Kruskal-Wallis test (Hollander et al. 2013) under R stats package. We also tested a possible link between isolate allelic richness and tree genetic background by grouping isolates by tree genetic background and performed Kruskal-Wallis test on A values obtained from those groups.

The genetic structure of sampled isolates was investigated using several complementary approaches. We first performed a PCA using ade4 R package (Dray et al. 2007). PCA allows a first visualization of the data that gives an idea of the global structure of the sample. In order to determine possible distinct genetic groups in our sample, we used STRUCTURE software (Pritchard et al. 2000). We used a 50 000 burnin period and 100 000 MCMC repetitions to explore group number (K) from K=1 to K=10 with 10 independent run for each K value. We used the admixture model with correlated allele frequencies and sample location information to increase group detection power (Hubisz et al. 2009). STRUCTURE results were interpreted using STRUCTURE HARVESTER website developed by Earl (2012) that applies Evanno et al. (2005) test to identify best K value. To quantify the effect of sampling location at various scales, we performed an AMOVA with poppr R package. We tested the effect of sampling location at three levels: first level was the isolate level; second level was plantation and finally third level with isolate grouped by country. Significance thresholds were calculated with 1000 permutations obtained with the ade4 R package. Isolation by distance phenomena was investigated with a classical Mantel test. We obtained physical distance matrix using QGIS version 2.6.1 and Nei (1978) genetic distance matrix using adegenet R package. Correlation between physical and genetic distances was assessed with the mantel test function implemented in ade4 R package using 1 000 permutations.

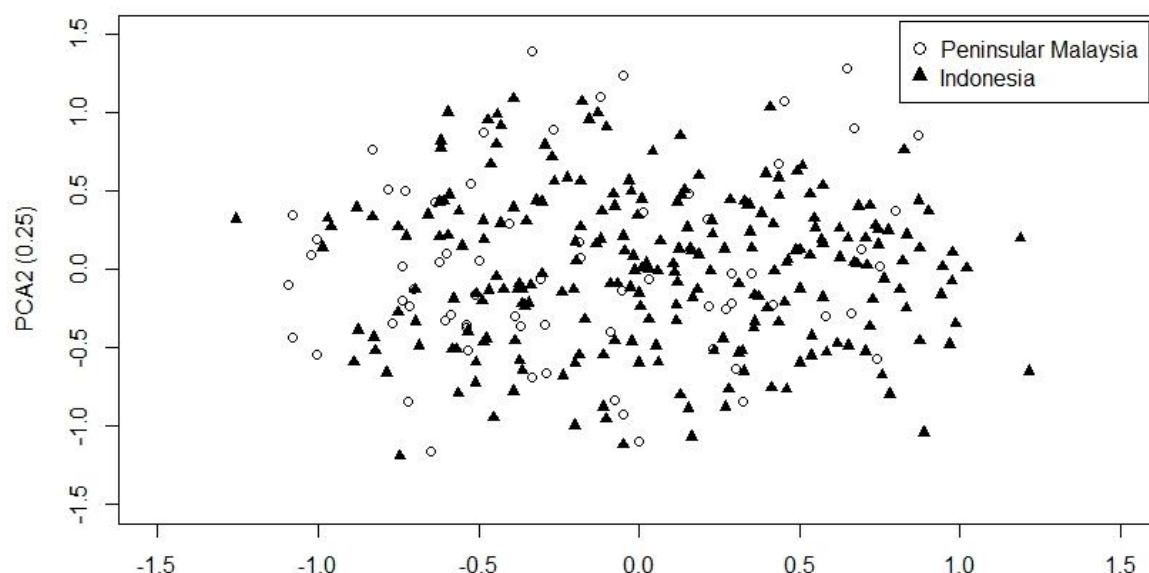
#### Demographic history of *Ganoderma boninense* in South-East Asia

We used the Approximate Bayesian Computation (ABC) approach as implemented in DIYABC software (Cornuet et al. 2014) to determine the demographic scenario that may have shaped *G. boninense* diversity pattern observed on SSR data. ABC is a recent flexible class of Monte-Carlo algorithms that performs model-based inference (Beaumont et al. 2002). ABC uses coalescence tool to generate thousands of simulated data having a similar configuration than observed one in terms of individual number, marker type and number and relies on summary statistics to summarize and compare simulated and observed data. Simulations are launched from a set of user-defined demographic scenarios with parameters sampled from prior distributions. DIYABC software provides a full set of tools that we used to simulate SSR data and to perform scenario choice, parameter evaluation and several steps of model validation.

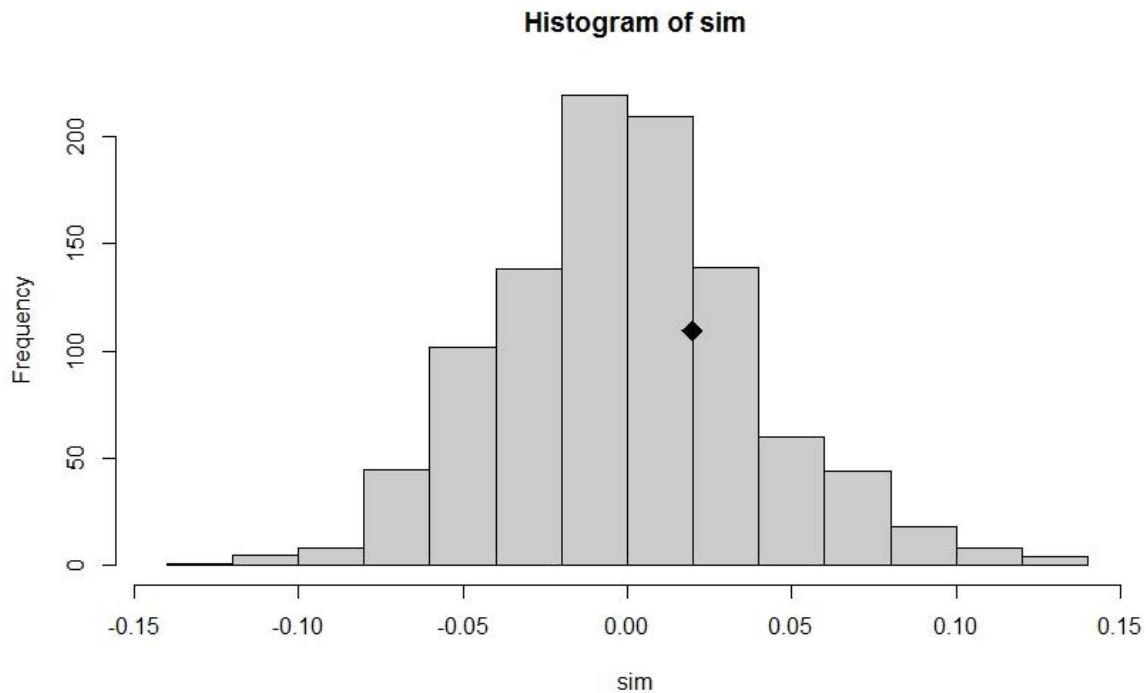
In our case, we used information from Structure and other genetic analysis to propose meaningful scenarios. We determined scenario posterior probability, parameter estimates and confidence in scenario choice and in parameter estimates using the statistical tools proposed in the DIYABC software. Scenario posterior probabilities were calculated both using a direct estimate (the proportion of point drawn from each scenario among the 500 simulated data points closest from the observed data) and a local logistic regression procedure proceeding on

the 10% closest simulated points and that, shortly saying, gives more weight to those closer to the observed data. Parameter posterior distributions were estimated using the 10% closest simulated points with a local regression procedure and was done using a logit transformation of the parameters. Confidence in scenario choice was evaluated by simulating an additional number of pseudo observed data (data obtained from resampling with parameters previously inferred from observed data) that are evaluated with the already existing reference table. Hence, for a given scenario, power is calculated as the number of pseudo observed data set drawn from this scenario and correctly assigned to it while alpha risk is calculated as the proportion of pseudo observed data set drawn from the other scenarios and wrongly assigned to the considered scenario. Confidence in parameter estimates is also evaluated using pseudo observed data points simulated with known parameters that are re-evaluated with the existing reference table. As a measure of parameter estimation error and bias, we calculated the average relative bias and the factor 2, i.e. the proportion of pseudo observed data set for which the point estimate is at least half and at most twice the true value (Cornuet et al. 2008). Additionally, and still using a DIYABC tool, we performed model checking in order to determine the adequacy to data of the chosen scenario, as advised by Cornuet et al. (2010). This model checking procedure consists in simulating pseudo-observed data set with parameters drawn from the posterior parameter distribution of the considered scenario and to position the summary statistics of the observed data in the summary statistic distribution of these pseudo observed data. The scenario is then considered suitable if the observed data summary statistics are included in the confidence interval drawn from pseudo observed data.

As we have poor knowledge on *Ganoderma boninense* biology we used as possible non-informative priors, meaning that priors were sampled from uniform distributions. We chose and studied scenario on the base of previous genetic analyses results and the few information we have on evolution of disease since the beginning of oil palm industrial culture.



**Figure 2.** Global genetic diversity observed by PCA analysis based on genotyped samples without missing data and using 11 microsatellite loci considered without null allele



**Figure 3.** Histogramme of simulated value distribution after 1 000 permutation compared to observed data. Simulated values are in gray and localization of observed data are represented by black square.

## Results

### Genetic analysis

Over the 357 samples collected and for which DNA was extracted, 311 exhibited amplification for all SSR markers. In order to perform population genetic analysis we searched for clones and removed 5 isolates from our dataset. Those 5 isolates considered as clones, were corresponding to independent fruiting bodies collected on different oil palm. Each pair of detected clones was corresponding to identical genotypes. All isolates pairs considered as clones came from carpophores collected on nearby trees (approximately 9 m between trees): 1 pair of clones came from Negeri Sembilan plantation, 2 pairs of clones came from Johor plantation, 1 pair of clones from Bangun Bandar plantation and 1 pair of clones came from Aek Loba plantation. We removed all loci with potential null allele as they could introduce bias in population genetic analysis. We detected 6 SSR markers with potential null allele. The 6 SSR markers likely bearing null alleles were loci 3d, 17b, 18a, 24d, 42d and 52d. Those loci were excluded from following analysis. Considering diversity at plantation level, the lowest  $H_e$  was obtained for samples collected in Aek Loba with a value of 0.521 and a standard deviation (SD) of 0.209 and the highest  $H_e$  value is obtained the samples collected in Bangun Bandar with a value of 0.677 and a SD of 0.084. Mean  $f$  allele number varied from 3.27 for Aek Loba plantation to 7.45 for Bangun Bandar plantation and allelic richness ( $R$ ) from 2.9 to 3.82 respectively for the same plantations. The Kruskal-Wallis test didn't show significant differences in allelic richness among plantations ( $P = 0.42$ ) or among regions ( $P = 0.32$ ). Kruskal-Wallis test revealed no significant difference between allelic richness obtained from isolates grouped by planting generation. Moreover, zooming on isolate from Sumatra plantation, allelic richness obtained from isolate grouped by tree genetic background reveal

no variation of allelic richness depending on tree genetic background with a Kruskal-Wallis P-value of 0.44.

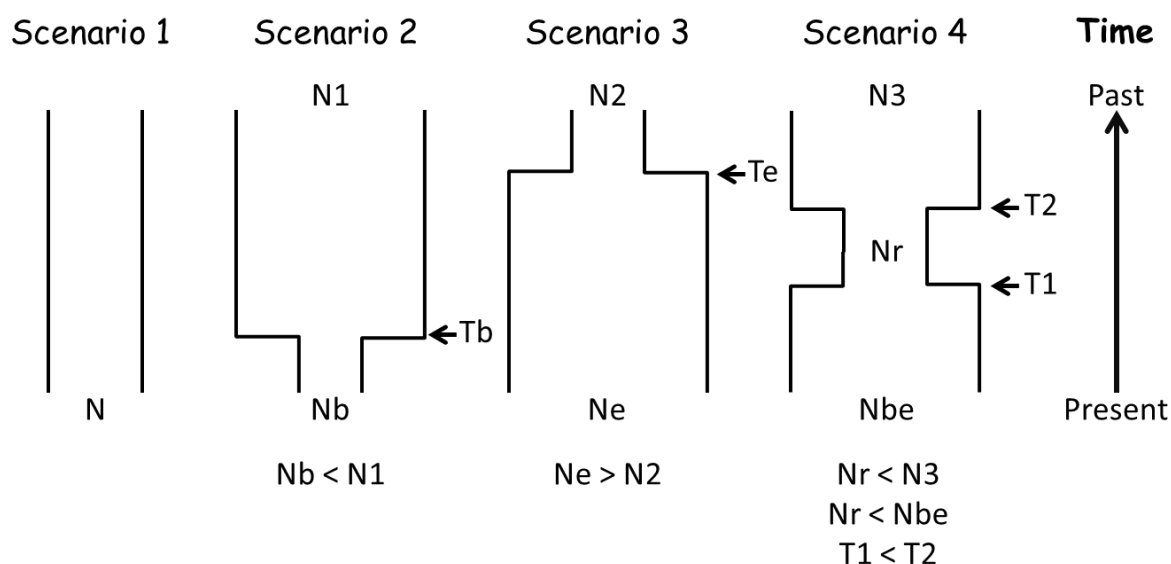
#### Population structuration of *Ganoderma boninense*

PCA analysis revealed no clear trend in the genetic structure of the data (Fig. 2). This result was confirmed by STRUCTURE which gave  $K=1$  as the most likely group number (Fig. S1). Those results are similar if we use marker set with null allele or without null allele, showing no significant influence of null allele on sample structuration.

As our isolates originate from a single population encompassing the two sampled regions with a maximum distance of 690 km between most distant collection points, we searched for a potential isolation by distance phenomena with a Mantel test. As indicated on Figure 3, this test revealed no significant difference between simulated data and observed data after permutations ( $p\text{-value}=0.274$ ), meaning no correlation between genetic and geographic distance. As no difference between samples from Peninsular Malaysia and Indonesia could be noticed, we searched for a smaller scale of sampling (plantation) which should highlight a population structuration by using hierarchical analyses of molecular variance (AMOVA) (Table 1). We observed a very low yet significant variation among region (1.59%,  $p\text{-value} < 0.001$ ) and among plantations (0.97%,  $p\text{-value} < 0.001$ ), and very high variation between samples (97.44%,  $p\text{-value} < 0.001$ ).

#### Demographic history of *Ganoderma boninense* in South-East Asia

In the absence of any kind of clear genetic structure, we decided to consider our samples as originating from a single panmictic population. In order to study demographic history of this population using ABC, we proposed four equi-probable scenarios: scenario n°1 depicting the



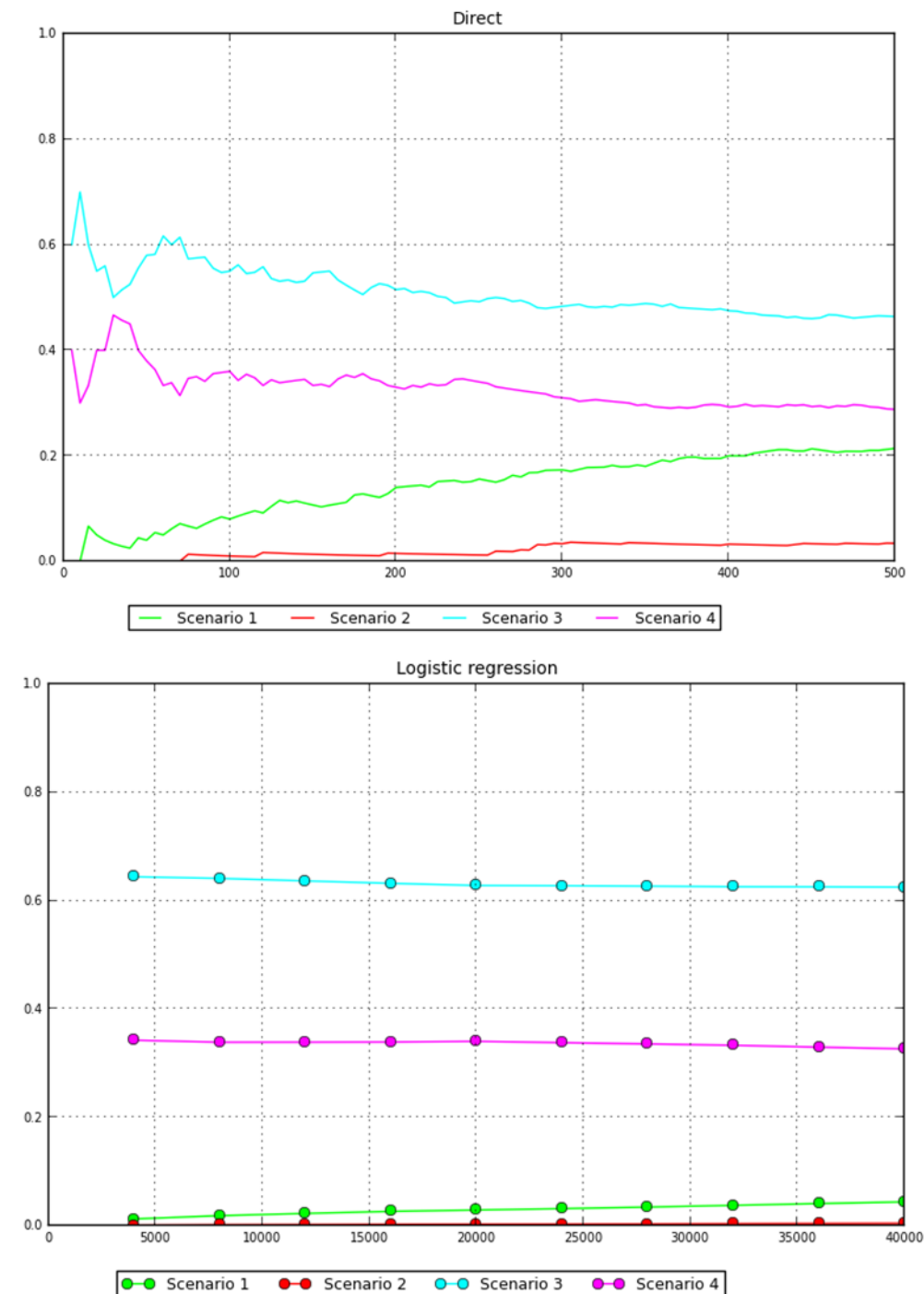
**Figure 4.** Demographic scenario for *Ganoderma boninense* in peninsular Malaysia and Sumatra using ABC approach as implemented in DYIABC 2.0.4 software by Cornuet et al. (2014). Scenario 1: evolution of a constant size  $N$  population. Scenario 2: population of size  $N1$  undergoes a bottleneck step  $Tb$  generation ago causing a significant reduction in the current population size  $Nb$ . Scenario 3: population of size  $Ne$  experiences a population expansion event from  $Te$  generation ago to actual population of size  $N2$ . Scenario 4: population of size  $N3$  undergo a bottleneck event  $T1$  generation ago causing a reduction in population size  $Nr$  followed by a population expansion  $T2$  generation ago to reach actual population size  $Nbe$ .

**Table 1.** Posteriors parameters estimation for scenario 3 and demographique expansion of scenario 4. Scenario 3: (Ne) Population size after expansion step, (Te) Time since beginning of expansion state in generation number, (N2) Population size before expansion, ( $\hat{\mu}_{mic}$ ) mean mutation rate, (pmic) number of repeat motifs added or removed from the microsatellite in each mutation step. Scenario 4: (Nbe) Population size after expansion step, (T1) Time since beginning of expansion state in generation number, (Nr) Population size before expansion.

Scenario	Parameter	mean	median	mode	q025	q975	Mean Bias	Fact2
Scenario 3	Ne	52700	49800	37000	12800	97200	0.196	0.914
	Te	3260	2370	769	81	9440	3.194	0.716
	N2	1780	1360	604	56.6	6500	0.629	0.718
	$\hat{\mu}_{mic}$	2,95E-04	2,19E-04	1,10E-04	1,04E-04	8,41E-04	0.270	0.838
	pmic	0.213	0.217	0.221	0.11	0.295	0.063	0.996
Scenario 4	Nbe	43200	37900	12000	5450	95900	0.246	0.9
	T1	2250	1810	552	91.7	6890	3.517	0.632
	Nr	800	569	218	55.9	3080	1.06	0.648
	$\hat{\mu}_{mic}$	4,49E-04	3,96E-04	1,43E-04	1,15E-04	9,70E-04	0.224	0.854
	pmic	0.202	0.201	0.172	0.113	0.294	0.077	0.998

evolution of a population of constant size  $N$  simulating no influence of oil palm as a new available host; scenario n°2 representing a possible bottleneck event due to a very recent host jump from coconut tree to oil palm tree which imply a drastic population size reduction, scenario n°3 testing for a demographic expansion step increasing the actual population size simulating accession of *G. boninense* to a new host considering that *G. boninense* have wide range of host without needing special adaptation from one host to another, and scenario n°4 presenting a more complex scenario were a first step of bottleneck is followed by a demographic expansion step simulating steps of a less recent host jump event. These four scenarios are represented in the figure 4 which also describes imposed parameter constraints. Those parameters were chosen as general and uninformative as possible since only limited information is available on *Ganoderma boninense* ecology and biology. The reference table was built with a total of 4 000 000 simulated data points, equally distributed among scenarios: 998 670 for scenario 1, 1 002 348 for scenario 2, 999 200 for scenario 3 and 999 782 for scenario 4. We used a subset of simulated points that were close enough to the observed data (Euclidian distance) and we recorded mean allele number, mean genetic diversity (Nei 1987), mean allele size variance and mean Garza–Williamson’s  $M$  (Garza et al. 2001) as sample statistics to estimate scenario probability, parameter confidence interval and parameter bias and error as well as to assess confidence in scenario choice. These parameters were chosen because a single population was considered in our simulations. Direct approach showed best support for scenario 3 and a rather good support for scenario 4 (Fig. 5), confirmed by high estimated posterior probability value and narrow confidence intervals, considering logistic regression approach and using 10% closet points: 0.6441 [0.6118,0.6765] for scenario 3 and 0.3424 [0.3100,0.3748] for scenario 4. In contrast, we observed very low posterior probabilities for scenario 1 and 2 (Fig. 5) (0.0125 [0.0000, 0.0292] for scenario 1 and 0.0009 [0.0000, 0.0178] for scenario 2). Considering these results we decide to keep only scenario 3 and 4 for further analysis. We performed a test on both scenario to estimate confidence in scenario choice and found that over 500 simulation scenario 3 was found with the highest posterior probability 320 times. In the light of this result, model checking was performed only for scenario 3. We constructed a new reference table for scenario 3 with same parameter priors in order to perform model checking analysis. We recorded only mean allele number and mean allele size variance as summary statistics for model choice and performed model checking with mean genetic diversity and mean Garza-Williamson’s  $M$ . Results have finally shown a strong support for scenario 3. Observing pre-evaluation of scenarios and prior distributions of summary statistics, scenario 3 show mean genetic diversity proportion (simulated<observed) of 0.45 and 0.24 for mean Garza-Williamson’s  $M$ . The main





**Figure 5.** Scenarios comparison based on posterior probabilities estimation. A show results obtained from direct approach and B from logistic regression method.

characteristics of parameter posterior distributions for scenarios 3 are presented in Table 1, along with the priors.

Even if the posterior probability of scenario 4 is lower than scenario 3 we decided to calculate estimation of posterior parameters for both scenarios as support obtained for scenario 4 may result from expansion step event, similar to scenario 3. Results of posterior parameters estimation for both expansion event are resumed in Table 1 and posterior distributions are drawn in Figure S1 (Supporting information). Regarding results obtained from expansion

step, scenario 3 present a starting population size ( $N_2$ ) of 1780 [56.6; 6 500] genotypes and scenario 4 a population size before expansion ( $N_r$ ) of 800 [55.9; 3 080]. Time of starting expansion is evaluated to a mean value of 3260 [81; 9 440] generations for scenario 3 and 2250 [91.7; 6 890] generations for scenario 4. Those generations times seem quite similar and gave us a mean value of 2 755 generations ago. Same results were obtained for the current population size with a mean value of 52 700 [12 800; 97 200] genotypes for scenario 3 and a mean of 43 200 [5 450; 95 900] genotypes for scenario 4.

## Discussion

### Diversity and dispersal of *Ganoderma boninense*

As expected, most of samples were successfully amplified by SSR markers without generating missing data or null allele for locus 7a, 11a, 19a, 22a, 33a, 6d, 11d, 20d, 37d, 47d and 51 d. This is a satisfactory result as markers detected without null allele corroborate those found in a previous study (Mercière et al. 2015). Moreover, two markers that were positive to null allele detection in previous study (Mercière et al. 2015) are now considered as free from null allele. This result should be due to an increase of information in our present data which increase the detection of marker without null allele. As the transfer of microsatellite markers from one species to another closely related one should generate null allele as suggested by Dutech et al. (2007), it is consistent with the fact that the 311 well amplified samples belongs to the species *Ganoderma boninense*. Our results show a rather high genetic diversity across peninsular Malaysia and Sumatra with global  $H_e$  value over the 11 SSR markers kept for analysis of 0.651 and a mean allele number of 9.27. Our results of  $H_e$  are quite higher than other results of root rot fungi like *Armillaria mellea* which were ranged between 0.103 and 0.600 depending on considered population (Prospero et al. 2008, Travadon et al. 2012 b). This result should support the hypothesis of an active sexual reproduction system. Pilotti et al. (2003) suggested in regards to a high diversity of mating type alleles that first mechanism of disease spreading is sexual reproduction by spore dispersal, supporting Miller (1995)'s results. The high diversity observed as well as the small number of duplicated genotypes found in our samples clearly indicates that sexual reproduction plays a key role in disease spreading, confirming Rees et al. (2012)'s results. Despite the fact that some palm could be infected by multiple genotypes (Miller 1995, Ariffin et al. 1996, Pilotti et al. 2003, Rees et al. 2012) and that we collected only one fruiting body at random for each palm tree, the small number of duplicated genotypes that we found (>2%) suggests that disease spreading by root-to-root infection is anecdotic. Duplicated genotypes were always collected on neighbor trees distant of about 9 meters while Oil Palm roots extend up to 30 meters. However, more precise estimation of clonal duplication events should be done in the future; as we sampled only one isolate per oil palm and that in industrial plantations careful sanitation is performed in order to avoid root-to-root contamination. Our results indicate a spreading pattern very different from *Armillaria mellea*, another model of perennial crop root rot basidiomycete, with a wide host spectrum and that mainly spreads by vegetative multiplication (Prospero et al. 2008, Baumgartner et al. 2010, Travadon et al. 2012 b). *G. boninense* dispersal mechanism seems more similar to *Heterobasium annosum*, a polyporales-like white rot fungi, which is responsible for important losses in conifers forests from northern temperate regions (Woodward et al. 1998) and has a predominantly sexual reproduction pattern.

### A single population

This high genetic diversity paired with an obvious spreading of the disease by basidiospores suggests a wide dispersal (like many others basidiomycetes) (Hallenberg et al. 2001) of *Ganoderma boninense* across South-East Asia. Pilotti et al. (2003), and more recently Rees et

al. (2012), showed a wide dispersal at the estate level but our PCA results hint a larger scale of spreading embracing both Sumatra and peninsular Malaysia. Results from Bayesian assignment test in STRUCTURE and AMOVA support that isolates from both regions originate from a single population, which indicates continual gene flows between those two regions in spite of the Malacca strait. This phenomenon of a large dispersal from 160 to 2 000 km by basidiospores was also noticed for the well-known white rot basidiomycetes *Armillaria mellea* in Northern America (Baumgartner et al. 2010, Travadon et al. 2012 b). Furthermore, the consideration of those samples as a single wide and homogenous population is reinforced by the absence of isolation by distance phenomena and by AMOVA results. Finally, the basidiospore dissemination can cover a total area of 200 000 km<sup>2</sup> with a maximum distance between two samples of 700 km. These basidiospores are most likely carried by wind (Rees et al. 2012), but some other factors (insects, birds, animals or human activity) could play a role in the basidiospore spreading. In fact, this airborne dispersal could be a shared characteristic between all phytopathogen polyporales.

The comparison of allelic richness level among isolates grouped by origins (plantation or region) show no significant between those groups. As each plantation/regions possess its own characteristics (like soil, climate ...) we can conclude that plantations conditions are not selective for any specific genotype of *G. boninense*. Same results were observed with isolates grouped by planting generation that shows no significant difference between those groups. This result could highlight the date of planting is not impacting the selection of specific genotypes which could be related to a planting year. It could also suggest that the effect of remaining inoculum in the soil from one generation to another, highlight by many authors (Hasan 1998, Breton et al. 2006, Rees et al. 2009), is not a major source of disease. This absence of correlation between factors and selection of specific genotype could also be biased by to high genetic diversity and the fact that microsatellite markers should be located in neutral genome region which should not be influenced by selection factors. Finally, isolates grouped by tree genetic background shows no significant difference between groups. This result suggests tree genetic background is not influencing the selection of specific subpopulations of *G. boninense*. It also could support the hypothesis of no specific interactions between oil palm and *G. boninense*.

#### Detection of a population expansion

We investigated demographic history of *G. boninense* using ABC approach and we compared 4 different models representing conceivable demographic history of *Ganoderma boninense* in Sumatra/Penninsular Malaysia. Since no evidence of population structure could be found, we considered that our samples belong to a single panmictic population. ABC analysis clearly supports a demographic expansion event in the *G. boninense* population from Sumatra/Penninsular Malaysia: the two scenarios proposing a demographic expansion event cumulate more than 98% of posterior probabilities. Among these two scenarios, scenario 3 (demographic expansion) has a higher posterior probability (64%) than scenario 4 (34%) that is characterized by a bottleneck followed by a demographic expansion. Such a bottleneck signature could be an artifact caused by various discrepancies between the real *G. boninense* population behavior and the Wright-Fisher assumption held in the ABC analysis. For example, high variance in reproductive success, that is a common feature in species with high fecundity, is known to generate fake bottleneck signature (Hoban et al. 2013). Estimation of posterior parameters indicates similar results in both scenario 3 and 4 for the expansion time with a mean value of 2 755 generations ago (between 2 250 generation ago for scenario 4 and 3 260 for scenario 3). Following Pilotti et al. (2003) who proposed a 4 to 7 year generation time, this expansion should have occurred between 12 800 and 22 820 years ago based on

scenario 3 expansion time. This estimation is not consistent with the hypothesis of *G. boninense* expansion along with oil palm plantation development since the beginning of the XXth century but rather indicates that expansion of *G. boninense* was related to some very ancient event like past climate changes. This period hence corresponds to the end of the last glacial period around 12 000 years ago (maximum of the extent glaciation around 22 000 years ago) which could have let some trace in actual *G. boninense* diversity like in many other species (Hewitt 2000).

The mean current population effective size estimated from both scenario seemed also similar in regard to the overlapping of their confidence interval values, and suggest a mean of 47 950 individuals. Such a large effective population size is consistent with the fact that *G. boninense* may have a wide host spectrum, like many white rot fungi (Baumgartner et al. 2010) (Woodward et al. 1998). However, estimations of population size and estimations of time expressed in generation number should be considered with caution as *G. boninense* has likely overlapping generations. Hence, in oil palm plantations, fruiting bodies can be active in spore release during several years (Breton, pers. com.).

## Conclusion

*G. boninense* isolates from Sumatra and peninsular Malaysia form a big and unique population. The global repartition of this population is not influenced neither by generation of planting nor planting material genetic background or physical obstacles. The population of *G. boninense* in those regions has undergone an expansion event which cannot be related to the beginning of the industrial expansion of oil palm culture but show some clues about relation between last glacial period and this expansion. To confirm this hypothesis it can be interesting to proceed with test in order to study time estimation between *G. boninense* generations.

## References

- Adamack, A. T. and B. Gruber (2014). PopGenReport: simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution* 5(4): 384-387.
- Agrios, G. N. (2005). *Plant Pathology*.
- Akbar, U., M. Kusnadi and M. Ollagnier (1971). Influence de la nature du materiel vegetal et de la nutrition minerale sur la pourriture seche du tronc du palmier a huile due a *Ganoderma*. *Oléagineux*.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). Basic local alignment search tool. *J Mol Biol* 215(3): 403-10.
- Anon (1958). Notes on current investigations and extension work, Octobre to December 1957. *Malay. agric.J.* 41: 51-59.
- Ariffin, D., A. Idris and A. Marzuki (1996). Spread of *Ganoderma boninense* and vegetative compatibility studies of a single field palm isolates. *Proceedings of the 1996 PORIM International Palm Oil Congress (Agriculture)*.
- Ariffin, D., A. Idris and G. Singh (2000). Status of *Ganoderma* in Oil 3 Palm. *Ganoderma diseases of perennial crops*: 49.
- Ariffin, D. I., A.S. (1993). Investigation on the control of *Ganoderma* with Dazomet. *PORIM Int. Palm oil Conf.*: 424-429.
- Arnaud-Haond, S. and K. Belkhir (2007). GENCLONE: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. *Molecular Ecology Notes* 7(1): 15-17.
- Arnaud-Haond, S., C. M. Duarte, F. Alberto and E. A. Serrao (2007). Standardizing methods to address clonality in population studies. *Mol Ecol* 16(24): 5115-39.

- Baldwin, B. G., M. J. Sanderson, J. M. Porter, M. F. Wojciechowski, C. S. Campbell and M. J. Donoghue (1995). The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden*: 247-277.
- Baumgartner, K., R. Travadon, J. Bruhn and S. E. Bergemann (2010). Contrasting Patterns of Genetic Diversity and Population Structure of *Armillaria mellea* sensu stricto in the Eastern and Western United States. *Phytopathology* 100(7): 708-718.
- Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste and F. Bonhomme (2004). genetix. Populations Genet Software Windows TM Univ Montp 4(05).
- Breton, F., Y. Hasan, S. Hariadi, Z. Lubis and H. De Franqueville (2006). Characterization of parameters for the development of an early screening test for basal stem rot tolerance in oil palm progenies. *Journal of Oil Palm Research*: 24-36.
- Cai, G., C. W. Leadbetter, M. F. Muehlbauer, T. J. Molnar and B. I. Hillman (2013). Genome-wide microsatellite identification in the fungus *Anisogramma anomala* using Illumina sequencing and genome assembly. *PLoS One* 8(11): e82408.
- Chapuis, M. P. and A. Estoup (2007). Microsatellite null alleles and estimation of population differentiation. *Mol Biol Evol* 24(3): 621-31.
- Chen, S., J. Xu, C. Liu, Y. Zhu, D. R. Nelson, S. Zhou, C. Li, L. Wang, X. Guo, Y. Sun, H. Luo, Y. Li, J. Song, B. Henrissat, A. Levasseur, J. Qian, J. Li, X. Luo, L. Shi, L. He, L. Xiang, X. Xu, Y. Niu, Q. Li, M. V. Han, H. Yan, J. Zhang, H. Chen, A. Lv, Z. Wang, M. Liu, D. C. Schwartz and C. Sun (2012). Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. *Nat Commun* 3: 913.
- Chung, G. F. (2011). Management of *Ganoderma* diseases in Oil palm plantations. *Planter* 87: 325-339
- Corley, R. H. V. and P. B. Tinker (2003). The oil palm. Oxford ; Malden, MA, Blackwell Science.
- Cornuet, J. M., P. Pudlo, J. Veyssier, A. Dehne-Garcia, M. Gautier, R. Leblois, J. M. Marin and A. Estoup (2014). DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*.
- Dabrowski, M. J., S. Bornelov, M. Kruczyk, N. Baltzer and J. Komorowski (2014). 'True' null allele detection in microsatellite loci: a comparison of methods, assessment of difficulties and survey of possible improvements. *Mol Ecol Resour* 15(3): 477-88.
- De Franqueville, H., H. Asmady, J. Jacquemard, Z. Hayun and T. Durand-Gasselin (2001). Indications on sources of oil palm (*Elaeis guineensis* Jacq.) genetic resistance and susceptibility to *Ganoderma* sp., the cause of basal stem rot. Cutting-edge technologies for sustained competitiveness: Proceedings of the 2001 PIPOC International Palm Oil Congress, Agriculture Conference, Kuala Lumpur, Malaysia, 20-22 August 2001., Malaysian Palm Oil Board (MPOB).
- Dennis, C. and J. Webster (1971). Antagonistic properties of species-groups of *Trichoderma*: I. Production of non-volatile antibiotics. *Transactions of the British Mycological Society* 57(1): 25-IN3.
- Doležel, J., P. Binarová and S. Lcretti (1989). Analysis of nuclear DNA content in plant cells by flow cytometry. *Biologia Plantarum* 31(2): 113-120.
- Dray, S. and A.-B. Dufour (2007). The ade4 package: implementing the duality diagram for ecologists. *Journal of statistical software* 22(4): 1-20.
- Drenth, A. (1994). Molecular genetic evidence for a new sexually reproducing population of *Phytophthora infestans* in Europe.
- Durand-Gasselin, T., H. Asmady, A. Flori, J. C. Jacquemard, Z. Hayun, F. Breton and H. de Franqueville (2005). Possible sources of genetic resistance in oil palm (*Elaeis guineensis*

Jacq.) to basal stem rot caused by *Ganoderma boninense*--prospects for future breeding. *Mycopathologia* 159(1): 93-100.

Dutech, C., J. Enjalbert, E. Fournier, F. Delmotte, B. Barres, J. Carlier, D. Tharreau and T. Giraud (2007). Challenges of microsatellite isolation in fungi. *Fungal Genet Biol* 44(10): 933-49.

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation genetics resources* 4(2): 359-361.

El Mousadik, A. and R. J. Petit (1996). High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theor Appl Genet* 92(7): 832-9.

Evanno, G., S. Regnaut and J. Goudet (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14(8): 2611-20.

Falcao, V. D. R., A. P. Tonon, M. C. Oliveira and P. Colepicolo (2008). RNA Isolation method for polysaccharide rich algae: agar producing *Gracilaria tenuistipitata* (Rhodophyta). *Journal of Applied Phycology* 20(1): 9-12.

Flood, J., P. D. Bridge and M. Holderness (2000 a). *Ganoderma* diseases of perennial crops. Wallingford, Oxon England ; New York, CABI.

Flood, J., Y. Hasan, P. Turner, E. O'Grady, P. Bridge and M. Holderness (2000 b). The spread of *Ganoderma* from infective sources in the field and its implications for management of the disease in oil palm. *Ganoderma* diseases of perennial crops: 101-112.

Flood, J., L. Keenan, S. Wayne and Y. Hasan (2005). Studies on oil palm trunks as sources of infection in the field. *Mycopathologia* 159(1): 101-107.

Friesen, T. L., E. H. Stukenbrock, Z. Liu, S. Meinhardt, H. Ling, J. D. Faris, J. B. Rasmussen, P. S. Solomon, B. A. McDonald and R. P. Oliver (2006). Emergence of a new disease as a result of interspecific virulence gene transfer. *Nature genetics* 38(8): 953-956.

Frison, E., C. Putter and M. Diekmann (1993). FAO/IBPGR technical guidelines for the safe movement of coconut germplasm, Bioversity International.

Fry, W. (2008). *Phytophthora infestans*: the plant (and R gene) destroyer. *Molecular plant pathology* 9(3): 385-402.

Garza, J. C. and E. G. Williamson (2001). Detection of reduction in population size using data from microsatellite loci. *Mol Ecol* 10(2): 305-18.

George, S. T. C., G.F.; Zakaria, K. (1996). Updated results (1990-1995) on trunk injection of fungicides for the control of *Ganoderma* basal stem rot. *PORIM Int. Palm oil Conf.*: 508-515.

Gurmit, S. (1991). *Ganoderma* - the scourge of oil palm in the coastal areas. *Planter, Kuala Lumpur* 67(421-444).

Hallenberg, N. and N. Kuffer (2001). Long-distance spore dispersal in wood-inhabiting Basidiomycetes. *Nordic Journal of Botany* 21(4): 431-436.

Hasan, Y. T., P.D. (1998). The comparative importance of different oil palm tissues as infection sources for basal stem rot in replantings. *Planter, Kuala Lumpur* 74: 119-135.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature* 405(6789): 907-913.

Ho, C. L. and Y. C. Tan (2014). Molecular defense response of oil palm to *Ganoderma* infection. *Phytochemistry*.

Ho, Y. and A. Nawawi (1986). Isolation, Growth and Sporophore Development of *Ganoderma boninense* from Oil Palm in Malaysia. *Pertanika* 9(1): 69-73.

Hoban, S. M., M. Mezzavilla, O. E. Gaggiotti, A. Benazzo, C. Van Oosterhout and G. Bertorelle (2013). High variance in reproductive success generates a false signature of a genetic bottleneck in populations of constant size: a simulation study. *BMC bioinformatics* 14(1): 309.

- Hollander, M., D. A. Wolfe and E. Chicken (2013). Nonparametric statistical methods, John Wiley & Sons.
- How, Y. and A. Nawawi (1985). *Ganoderma boninense* Pat. from basal stem rot of oil palm (*Elaeis guineensis*) in Peninsular Malaysia. *Pertanika* 8(3): 425-428.
- Hubisz, M. J., D. Falush, M. Stephens and J. K. Pritchard (2009). Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour* 9(5): 1322-32.
- Hyman, E. D. (1988). A new method of sequencing DNA. *Analytical biochemistry* 174(2): 423-436.
- Idris, A., D. Ariffin, T. Watt and T. Swinburne (2001). Distribution of species of *Ganoderma* basal stem rot of oil palms in relation to the environmental conditions in peninsular Malaysia. Cutting-edge technologies for sustained competitiveness: Proceedings of the 2001 PIPOC International Palm Oil Congress, Agriculture Conference, Kuala Lumpur, Malaysia, 20-22 August 2001., Malaysian Palm Oil Board (MPOB).
- Idris, A., R. Arifurrahman and A. Kushairi (2010). Hexaconazole as a preventive treatment for managing *Ganoderma* in oil palm. MPOB TS Information Series 75.
- Idris, A., S. Ismail, D. Ariffin and H. Ahmad (2004). Prolonging the productive life of *Ganoderma* infected palms with hexaconazole. MPOB Information Series(221).
- Idris, A., A. Kushairi, S. Ismail and D. Ariffin (2004). Selection for partial resistance in oil palm progenies to *Ganoderma* basal stem rot. *J Oil Palm Res* 16(2): 12-18.
- Idris, A. and A. Rafidah (2008). Polyclonal antibody for detection of *Ganoderma*. MPOB Information Series, MPOB TT(405).
- Idris, A., S. Rajinder, A. Madihah and B. W. Mohd (2010). Multiplex PCR-DNA kit for early detection and identification of *Ganoderma* species in oil palm. MPOB Inf Ser, MPOB TS(73).
- Idris, A. S. M., S.M. (2012). PROLONGING THE PRODUCTIVE LIFE OF *Ganoderma*-INFECTED OIL PALM WITH DAZOMET. MPOB Information Series 616.
- Idris, A. S. Y., M.; Hayakawa, S.; Basri, M.W.; Noorhasimah, I.; Ariffin, D. (2003 ). PCR technique for detection of *Ganoderma*. MPOB TT 188
- Jacquemard, J.-C. (2011). Le palmier à huile, Editions Quae.
- Jarne, P. and P. J. L. Lagoda (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution* 11(10): 424-429.
- Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24(11): 1403-5.
- Kalinowski, S. T. and M. L. Taper (2006). Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. *Conservation Genetics* 7(6): 991-995.
- Larter, L. N. H. (1956). Notes on current investigations, April to June. *Malay. agric.J.* 34: 220-227.
- Lepoivre, P. (2003). *Phytopathologie*. 1 éd. Bruxelles: De Boeck Université.
- Li, H. and R. Durbin (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25(14): 1754-1760.
- Martinez, A., J. Barrasa, M. Martinez, G. Almendros, M. Blanco and A. González (1995). *Ganoderma australe*: a fungus responsible for extensive delignification of some Austral hardwoods. *Ganoderma*. Systematics, phytopathology and pharmacology. Taipei: National Taiwan University. p: 67-77.
- McDonald, B. A. (1997). The population genetics of fungi: tools and techniques. *Phytopathology* 87(4): 448-53.
- McDonald, B. A. and C. Linde (2002). The population genetics of plant pathogens and breeding strategies for durable resistance. *Euphytica* 124(2): 163-180.
- Miller, R. N. G. (1995). The characterization of *Ganoderma* populations in oil palm cropping systems, University of Reading.



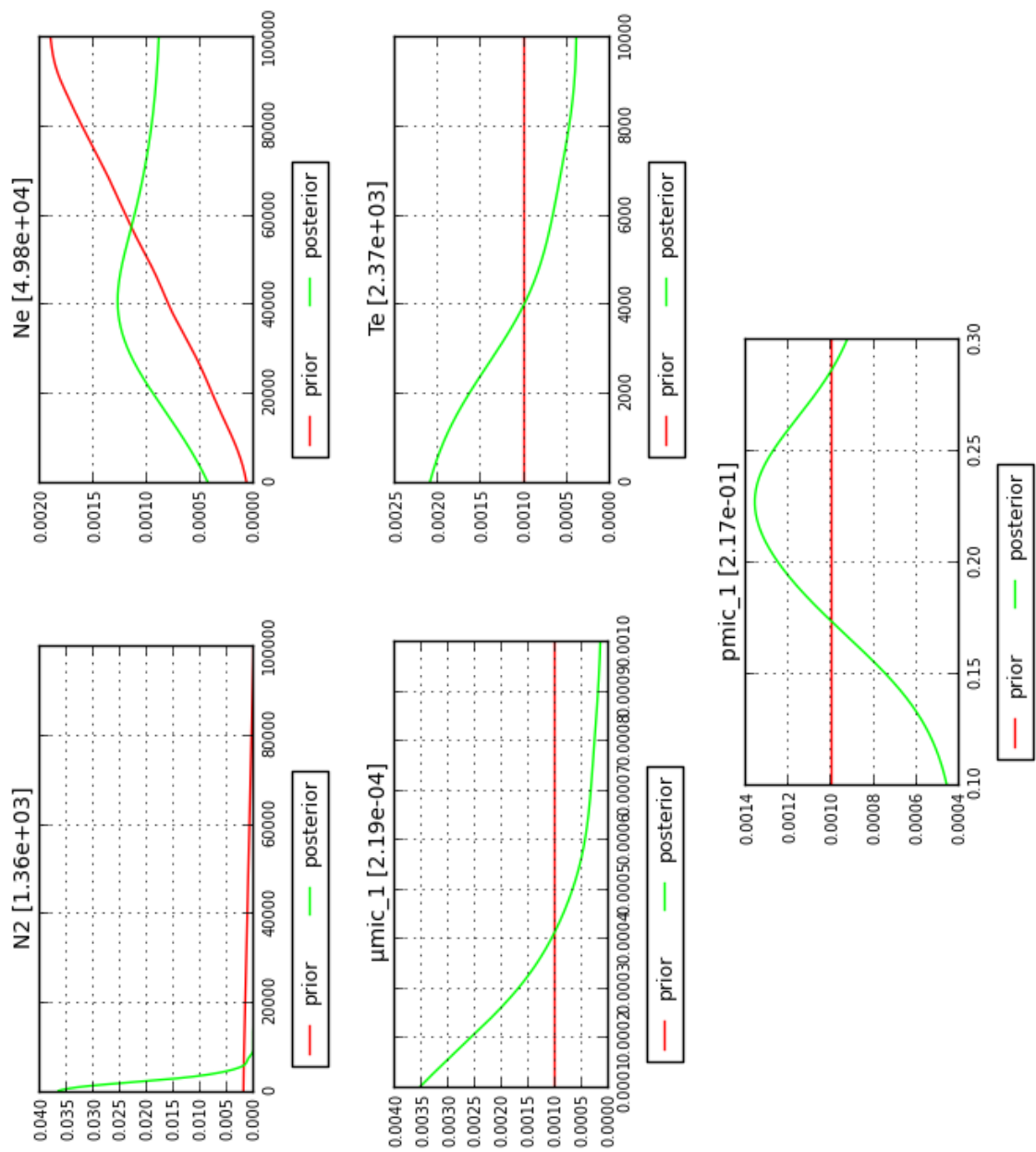
- Miller, R. N. G., M. Holderness, P. D. Bridge, G. F. Chung and M. H. Zakaria (1999). Genetic diversity of *Ganoderma* in oil palm plantings. *Plant Pathology* 48(5): 595-603.
- Moncalvo, J. (2000). Systematics of *Ganoderma* 2. *Ganoderma* diseases of perennial crops: 23.
- Moncalvo, J. M., H. H. Wang and R. S. Hseu (1995 b). Phylogenetic-Relationships in *Ganoderma* Inferred from the Internal Transcribed Spacers and 25s Ribosomal DNA-Sequences. *Mycologia* 87(2): 223-238.
- Nei, M. (1987). Molecular evolutionary genetics, Columbia university press.
- Paetkau, D. and C. Strobeck (1995). The molecular basis and evolutionary history of a microsatellite null allele in bears. *Mol Ecol* 4(4): 519-20.
- Pilotti, C. A., F. R. Sanderson and E. A. B. Aitken (2002). Sexuality and interactions of monokaryotic and dikaryotic mycelia of *Ganoderma boninense*. *Mycological Research* 106: 1315-1322.
- Pilotti, C. A., F. R. Sanderson and E. A. B. Aitken (2003). Genetic structure of a population of *Ganoderma boninense* on oil palm. *Plant Pathology* 52(4): 455-463.
- Presti, F. T. and A. P. Wasko (2014). A Review of Microsatellite Markers and their Application on Genetic Diversity Studies in Parrots. *Open Journal of Genetics* 2014.
- Pritchard, J. K., M. Stephens and P. Donnelly (2000). Inference of population structure using multilocus genotype data. *Genetics* 155(2): 945-59.
- Prospero, S., B. Lung-Escarmant and C. Dutech (2008). Genetic structure of an expanding Armillaria root rot fungus (*Armillaria ostoyae*) population in a managed pine forest in southwestern France. *Mol Ecol* 17(14): 3366-78.
- Qian, J., H. Xu, J. Song, J. Xu, Y. Zhu and S. Chen (2013). Genome-wide analysis of simple sequence repeats in the model medicinal mushroom *Ganoderma lucidum*. *Gene* 512(2): 331-6.
- Raes, N., C. H. Cannon, R. J. Hijmans, T. Piessens, L. G. Saw, P. C. van Welzen and J. W. F. Slik (2014). Historical distribution of Sundaland's Dipterocarp rainforests at Quaternary glacial maxima. *Proceedings of the National Academy of Sciences* 111(47): 16790-16795.
- Rees, R. W., J. Flood, Y. Hasan and R. M. Cooper (2007). Effects of inoculum potential, shading and soil temperature on root infection of oil palm seedlings by the basal stem rot pathogen *Ganoderma boninense*. *Plant Pathology* 56(5): 862-870.
- Rees, R. W., J. Flood, Y. Hasan, U. Potter and R. M. Cooper (2009). Basal stem rot of oil palm (*Elaeis guineensis*); mode of root infection and lower stem invasion by *Ganoderma boninense*. *Plant Pathology* 58(5): 982-989.
- Rees, R. W., J. Flood, Y. Hasan, M. A. Wills and R. M. Cooper (2012). *Ganoderma boninense* basidiospores in oil palm plantations: evaluation of their possible role in stem rots of *Elaeis guineensis*. *Plant Pathology* 61(3): 567-578.
- Robinson, M. D., D. J. McCarthy and G. K. Smyth (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26(1): 139-140.
- Roughgarden, J. (1979). Theory of population genetics and evolutionary ecology: an introduction.
- Ryvarden, L. (1985). Type studies in the Polyporaceae. 17. Species described by WA Murrill. *Mycotaxon*.
- Saïdi, M. N., R. Gargouri-Bouzid, M. Rayanni and N. Drira (2009). Optimization of RNA isolation from brittle leaf disease affected date palm leaves and construction of a subtractive cDNA library. *Molecular biotechnology* 41(1): 63-68.
- Sariah, M., H. Zakaria, J. Flood, P. Bridge and M. Holderness (2000). The use of soil amendments for the control of basal stem rot of oil-palm seedlings. *Ganoderma* diseases of perennial crops: 89-99.

- Schoebel, C. N., S. Brodbeck, D. Buehler, C. Cornejo, J. Gajurel, H. Hartikainen, D. Keller, M. Leys, S. Ricanova, G. Segelbacher, S. Werth and D. Csencsics (2013 b). Lessons learned from microsatellite development for nonmodel organisms using 454 pyrosequencing. *J Evol Biol* 26(3): 600-11.
- Schoebel, C. N., E. Jung and S. Prospero (2013 a). Development of new polymorphic microsatellite markers for three closely related plant-pathogenic *Phytophthora* species using 454-pyrosequencing and their potential applications. *Phytopathology* 103(10): 1020-7.
- Sharples, A. (1928). Palm diseases in Malaya. *Malayan Agricultural Journal* 16(9-10): 313-360.
- Steyaert, R. L. (1967). Les *Ganoderma* palmicoles. *Bulletin du Jardin Botanique Nationale de Belgique*(37): 465-492.
- Susanto, A., P. S. Sudharto and R. Y. Purba (2005). Enhancing biological control of basal stem rot disease (*Ganoderma boninense*) in oil palm plantations. *Mycopathologia* 159(1): 153-7.
- Tan, Y. C., K. A. Yeoh, M. Y. Wong and C. L. Ho (2013). Expression profiles of putative defence-related proteins in oil palm (*Elaeis guineensis*) colonized by *Ganoderma boninense*. *Journal of Plant Physiology* 170(16): 1455-1460.
- Tee, S. S., Y. C. Tan, F. Abdullah, M. Ong-Abdullah and C. L. Ho (2013). Transcriptome of oil palm (*Elaeis guineensis* Jacq.) roots treated with *Ganoderma boninense*. *Tree Genetics & Genomes* 9(2): 377-386.
- Thomas, F., T. Lefèvre and M. Raymond (2010). *Biologie évolutive*.
- Torres, G., G. Sarria, F. Varon, M. Coffey, M. Elliott and G. Martinez (2010). First report of bud rot caused by *Phytophthora palmivora* on African oil palm in Colombia. *Plant Disease* 94(9): 1163-1163.
- Travadon, R., M. E. Smith, P. Fujiyoshi, G. W. Douhan, D. M. Rizzo and K. Baumgartner (2012 b). Inferring dispersal patterns of the generalist root fungus *Armillaria mellea*. *New Phytologist* 193(4): 959-969.
- Tronsmo, A. and C. Dennis (1977). The use of *Trichoderma* species to control strawberry fruit rots. *Netherlands Journal of Plant Pathology* 83(1): 449-455.
- Turner, P. (1965 a). The incidence of *Ganoderma* disease of oil palms in Malaya and its relation to previous crop. *Annals of Applied Biology* 55(3): 417-423.
- Turner, P. (1965 b). Infection of oil palms by *Ganoderma*, AMER PHYTOPATHOLOGICAL SOC 3340 PILOT KNOB ROAD, ST PAUL, MN 55121. 55: 937-&.
- Turner, P. D. (1965 c). Oil palms and *Ganoderma* - is there a problem? *Planter, Kuala Lumpur* 41: 187-189.
- Turner, P. D. (1981). Oil palm diseases and disorders. *Oil palm diseases and disorders*.
- Turner, P. D. and R. Gillbanks (1974). Oil palm cultivation and management. *Oil palm cultivation and management*.
- Utomo, C., S. Werner, F. Niepold and H. B. Deising (2005). Identification of *Ganoderma*, the causal agent of basal stem rot disease in oil palm using a molecular method. *Mycopathologia* 159(1): 159-70.
- Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills and P. Shipley (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4(3): 535-538.
- Varghese, G., P. Chew and J. Lim (1976). Biology and chemically assisted biological control of *Ganoderma*. *Proceedings, International Rubber Conference, Vol. III., Rubber Research Institute of Malaysia*.
- Wakefield, E. (1920). Diseases of the oil palm in West Africa. *Bulletin of Miscellaneous Information (Royal Botanic Gardens, Kew)*: 306-308.

- Woodward, S., J. Stenlid, R. Karjalainen and A. Hüttermann (1998). *Heterobasidion annosum*: biology, ecology, impact and control, Cab International.
- Yow, S. and J. Nasir (2001). Replanting policies and strategies in Golden Hope. Cutting-edge technologies for sustained competitiveness: Proceedings of the 2001 PIPOC International Palm Oil Congress, Agriculture Conference, Kuala Lumpur, Malaysia, 20-22 August 2001., Malaysian Palm Oil Board (MPOB).
- Zane, L., L. Bargelloni and T. Patarnello (2002). Strategies for microsatellite isolation: a review. *Mol Ecol* 11(1): 1-16.

Supplementary information

SI Figure 1.-Posterior distribution of estimate parameters for scenario 3



**SI Table 1.** Accessory informations from genotyped samples without any missing data

Sample ID	Region	Plantation	Block	Planting Year	Planting Generation	Tree Symptomes	Tree Genetic Background
70	Peninsular Malaysia	Johor	J2	1989	Second	important	Unknow
71			J2	1989	Second	important	Unknow
72			J2	1989	Second	dead	Unknow
73			J2	1989	Second	important	Unknow
74			J2	1989	Second	medium	Unknow
75			J2	1989	Second	healthy	Unknow
76			J2	1989	Second	important	Unknow
77			J2	1989	Second	dead	Unknow
78			J2	1989	Second	important	Unknow
79			J3	Unknow	Second	important	Unknow
80			J3	Unknow	Second	light	Unknow
81			J3	Unknow	Second	light	Unknow
82			J3	Unknow	Second	healthy	Unknow
83			J3	Unknow	Second	dead	Unknow
84			J3	Unknow	Second	dead	Unknow
85			J3	Unknow	Second	important	Unknow
86			J3	Unknow	Second	medium	Unknow
87			J3	Unknow	Second	dead	Unknow
88,1			J3	Unknow	Second	important	Unknow
88,2			J3	Unknow	Second	important	Unknow
88,3			J3	Unknow	Second	important	Unknow
88,4			J3	Unknow	Second	important	Unknow
89			J3	Unknow	Second	light	Unknow
90			J3	Unknow	Second	medium	Unknow
91			J3	Unknow	Second	light	Unknow
92			J2	1989	Second	important	Unknow
93			J2	1989	Second	medium	Unknow
94			J2	1989	Second	light	Unknow
95			J4	Unknow	Second	dead	Unknow
96			J4	Unknow	Second	dead	Unknow
97			J4	Unknow	Second	medium	Unknow
98			J4	Unknow	Second	light	Unknow
99			J4	Unknow	Second	medium	Unknow
100			J4	Unknow	Second	medium	Unknow
101			J4	Unknow	Second	medium	Unknow
103			J4	Unknow	Second	medium	Unknow
104			J4	Unknow	Second	important	Unknow
105			J4	Unknow	Second	important	Unknow
106			J4	Unknow	Second	dead	Unknow
107			J4	Unknow	Second	dead	Unknow
108			J4	Unknow	Second	dead	Unknow
109			J4	Unknow	Second	medium	Unknow
110			J4	Unknow	Second	dead	Unknow
111			J4	Unknow	Second	important	Unknow
112			J3	Unknow	Second	important	Unknow
113		Negeri Sembilan	b3	Unknow	Unknow	healthy	Unknow
114			b3	Unknow	Unknow	medium	Unknow

Sample ID	Region	Plantation	Block	Planting Year	Planting Generation	Tree Symptomes	Tree Genetic Background
115	Peninsular Malaysia	Negeri Sembilan	b3	Unknow	Unknow	medium	Unknow
116			b3	Unknow	Unknow	light	Unknow
117			b3	Unknow	Unknow	medium	Unknow
118			b3	Unknow	Unknow	healthy	Unknow
119			b3	Unknow	Unknow	light	Unknow
120			b3	Unknow	Unknow	important	Unknow
121			b3	Unknow	Unknow	important	Unknow
122			b3	Unknow	Unknow	important	Unknow
123			b3	Unknow	Unknow	light	Unknow
124			b3	Unknow	Unknow	light	Unknow
125			b3	Unknow	Unknow	important	Unknow
126			b3	Unknow	Unknow	important	Unknow
127			b3	Unknow	Unknow	medium	Unknow
128			b3	Unknow	Unknow	important	Unknow
129			b3	Unknow	Unknow	medium	Unknow
130			b3	Unknow	Unknow	light	Unknow
131			b3	Unknow	Unknow	medium	Unknow
133			b3	Unknow	Unknow	medium	Unknow
134			b3	Unknow	Unknow	dead	Unknow
135			b3	Unknow	Unknow	healthy	Unknow
136			b3	Unknow	Unknow	important	Unknow
137			b3	Unknow	Unknow	medium	Unknow
138			b4	Unknow	Unknow	medium	Unknow
139			b4	Unknow	Unknow	light	Unknow
140			b4	Unknow	Unknow	light	Unknow
141			b4	Unknow	Unknow	important	Unknow
142			b4	Unknow	Unknow	medium	Unknow
143			b4	Unknow	Unknow	light	Unknow
144			b4	Unknow	Unknow	important	Unknow
145			b4	Unknow	Unknow	medium	Unknow
146			b4	Unknow	Unknow	medium	Unknow
147			b4	Unknow	Unknow	medium	Unknow
148			b4	Unknow	Unknow	medium	Unknow
150			b4	Unknow	Unknow	important	Unknow
151			b4	Unknow	Unknow	light	Unknow
152			b4	Unknow	Unknow	medium	Unknow
153			b4	Unknow	Unknow	light	Unknow
154			b4	Unknow	Unknow	medium	Unknow
156			b4	Unknow	Unknow	light	Unknow
157			b4	Unknow	Unknow	medium	Unknow
158			b4	Unknow	Unknow	light	Unknow
159			b4	Unknow	Unknow	medium	Unknow
160			b4	Unknow	Unknow	healthy	Unknow
164			b4	Unknow	Unknow	light	Unknow
165			b4	Unknow	Unknow	important	Unknow
200	Sumatra	Bangun Bandar	BB55	2001	Second	important	IL-6305
201			BB55	2001	Second	light	IL-6305

Sample ID	Region	Plantation	Block	Planting Year	Planting Generation	Tree Symptomes	Tree Genetic Background
202	Sumatra	Bangun Ban dar	BB55	2001	Second	medium	IL-6305
203			BB55	2001	Second	important	IL-6305
204			BB55	2001	Second	dead	IL-6305
206			BB55	2001	Second	important	IL-6305
207			BB55	2001	Second	important	IL-6305
208			BB55	2001	Second	medium	IL-6305
209			BB55	2001	Second	medium	IL-6305
210			BB61	1996	Second	important	MIXED
211			BB61	1996	Second	medium	MIXED
212			BB61	1996	Second	medium	MIXED
213			BB61	1996	Second	dead	MIXED
214			BB61	1996	Second	important	MIXED
215			BB61	1996	Second	important	MIXED
216			BB61	1996	Second	healthy	MIXED
217			BB61	1996	Second	medium	MIXED
218			BB60	1999	First	medium	Perc.PSBB
219			BB60	1999	First	healthy	Perc.PSBB
220			BB60	1999	First	important	Perc.PSBB
221			BB60	1999	First	important	Perc.PSBB
222			BB60	1999	First	medium	Perc.PSBB
223			BB60	1999	First	light	Perc.PSBB
224			BB60	1999	First	important	Perc.PSBB
226			BB57a	1991	Second	medium	MIXED
227			BB57a	1991	Second	important	MIXED
228			BB57a	1991	Second	dead	MIXED
229			BB57a	1991	Second	medium	MIXED
230			BB57a	1991	Second	medium	MIXED
231			BB57a	1991	Second	dead	MIXED
232			BB57a	1991	Second	light	MIXED
234			BB57b	2003	First	important	S-6803
237			BB57b	1997	First	important	S-6803
238			BB57b	1997	First	light	S-6803
239			BB57b	1997	First	important	S-6803
240			BB58	1999	Second	medium	IL-6305
241			BB58	1999	Second	important	IL-6305
242			BB58	1999	Second	important	IL-6305
243			BB58	1999	Second	medium	IL-6305
244			BB58	1999	Second	medium	IL-6305
245			BB58	1999	Second	medium	IL-6305
246			BB58	1999	Second	medium	IL-6305
247			BB42	2000	Second	important	IL-6305
248			BB42	2000	Second	important	IL-6305
250			BB42	2000	Second	light	IL-6305
251			BB42	2000	Second	light	IL-6305
252			BB42	2000	Second	important	IL-6305
253			BB42	2000	Second	dead	IL-6305
254			BB42	2000	Second	important	IL-6305
255			BB42	2000	Second	dead	IL-6305



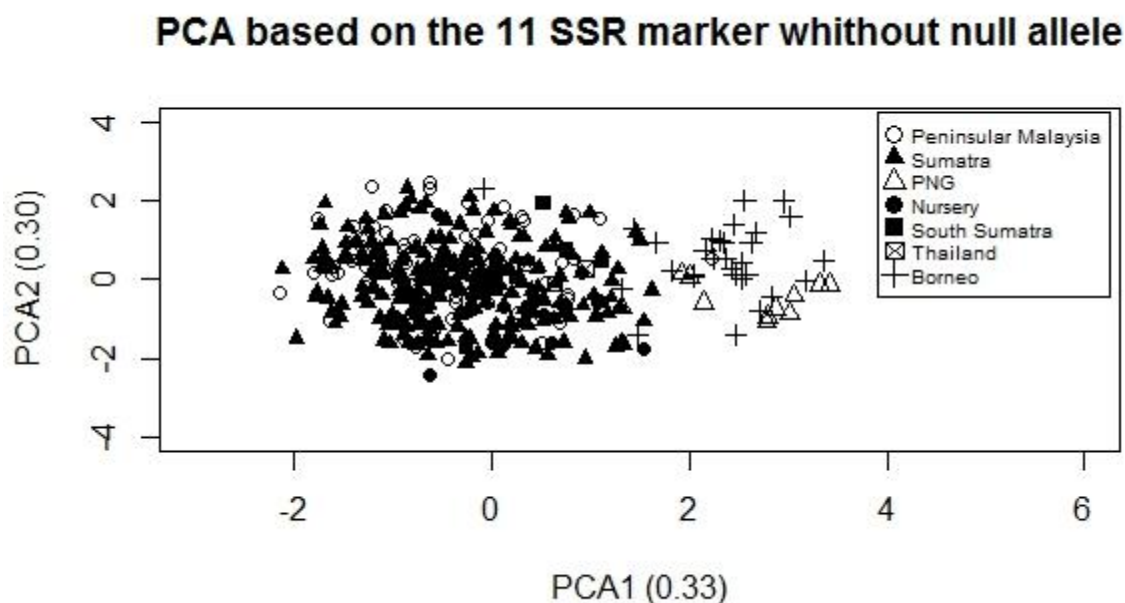
Sample ID	Region	Plantation	Block	Planting Year	Planting Generation	Tree Symptomes	Tree Genetic Background
256	Sumatra	Bangun Bandar	BB39	1998	Second	light	IL-5005
257			BB39	1998	Second	dead	IL-5005
258			BB39	1998	Second	important	IL-5005
259			BB39	1998	Second	important	IL-5005
263			BB25a	1989	Third	light	Percobaan/Klonal/IL-4905
264			BB25a	1989	Third	medium	Percobaan/Klonal/IL-4905
265			BB25a	1989	Third	important	Percobaan/Klonal/IL-4905
269			BB25b	1989	Third	light	Percobaan/Klonal/IL-4905
270			BB19a	1990	Third	light	IL-4905
271			BB19a	1990	Third	medium	IL-4905
272			BB19a	1990	Third	medium	IL-4905
273			BB19a	1990	Third	dead	IL-4905
274			BB19b	1990	Third	healthy	IL-4905
275			BB19b	1990	Third	light	IL-4905
276			BB19b	1990	Third	important	IL-4905
277			BB19b	1990	Third	medium	IL-4905
278			BB19b	1990	Third	important	IL-4905
279			BB19b	1990	Third	dead	IL-4905
280			BB19b	1990	Third	important	IL-4905
281			BB19b	1990	Third	important	IL-4905
282			BB07	2001	Second	light	IL-5005
283			BB07	2001	Second	important	IL-5005
284			BB07	2001	Second	important	IL-5005
285			BB07	2001	Second	important	IL-5005
286			BB07	2001	Second	medium	IL-5005
287			BB07	2001	Second	medium	IL-5005
288			BB07	2001	Second	important	IL-5005
289			BB11	1990	Second	important	IL-4906
290			BB11	1990	Second	medium	IL-4906
291			BB11	1990	Second	light	IL-4906
292			BB11	1990	Second	important	IL-4906
293			BB11	1990	Second	light	IL-4906
294			BB11	1990	Second	dead	IL-4906
295			BB11	1990	Second	light	IL-4906
296			BB11	1990	Second	medium	IL-4906
298			BB11	1990	Second	light	IL-4906
299			BB16	1991	Third	medium	IL-4905
300			BB16	1991	Third	light	IL-4905
301			BB16	1991	Third	healthy	IL-4905
302			BB16	1991	Third	medium	IL-4905
303			BB16	1991	Third	healthy	IL-4905
304			BB16	1991	Third	light	IL-4905
305			BB16	1991	Third	important	IL-4905
306			BB31	1992	Second	medium	IL-6305

Sample ID	Region	Plantation	Block	Planting Year	Planting Generation	Tree Symptomes	Tree Genetic Background
307	Sumatra	Bangun Ban dar	BB 31	1992	Second	healthy	IL-6305
308			BB 31	1992	Second	important	IL-6305
309			BB 31	1992	Second	important	IL-6305
310			BB 31	1992	Second	medium	IL-6305
312		Mata Pau	MP07	1999	Third	medium	IL-6305
313			MP07	1999	Third	healthy	IL-6305
314			MP07	1999	Third	medium	IL-6305
315			MP07	1999	Third	medium	IL-6305
316			MP07	1999	Third	important	IL-6305
317			MP07	1999	Third	medium	IL-6305
318			MP13	1989	First	important	IN-0707
319			MP13	1989	First	medium	IN-0707
320			MP13	1989	First	medium	IN-0707
321			MP13	1989	First	dead	IN-0707
322			MP13	1989	First	important	IN-0707
323			MP13	1989	First	medium	IN-0707
324			MP13	1989	First	light	IN-0707
325			MP25	2000	Fourth	important	Percobaan
326			MP25	2000	Fourth	dead	Percobaan
327			MP25	2000	Fourth	dead	Percobaan
328			MP25	2000	Fourth	light	Percobaan
329			MP25	2000	Fourth	healthy	Percobaan
330			MP25	2000	Fourth	healthy	Percobaan
331			MP25	2000	Fourth	healthy	Percobaan
332			MP25	2000	Fourth	light	Percobaan
333			MP25	2000	Fourth	healthy	Percobaan
334			MP30	1999	Fourth	light	IL-5105
335			MP30	1999	Fourth	important	IL-5105
336			MP30	1999	Fourth	dead	IL-5105
337			MP30	1999	Fourth	medium	IL-5105
338			MP30	1999	Fourth	light	IL-5105
339			MP30	1999	Fourth	healthy	IL-5105
340			MP35	2001	Fourth	dead	IL-6305
341			MP35	2001	Fourth	important	IL-6305
342			MP35	2001	Fourth	important	IL-6305
343			MP35	2001	Fourth	medium	IL-6305
345			MP35	2001	Fourth	dead	IL-6305
346			MP35	2001	Fourth	medium	IL-6305
347			MP39	1993	Fourth	medium	IL-6305
348			MP40	1993	Fourth	important	Klon al P erc. PSBB
349			MP40	1993	Fourth	important	Klon al P erc. PSBB
350			MP40	1993	Fourth	dead	Klon al P erc. PSBB
351			MP40	1993	Fourth	important	Klon al P erc. PSBB
352			MP40	1993	Fourth	important	Klon al P erc. PSBB
353			MP40	1993	Fourth	medium	Klon al P erc. PSBB
355			MP40	1993	Fourth	light	Klon al P erc. PSBB
357			MP43	1995	Fourth	healthy	IL-5005
360			MP43	1995	Fourth	light	IL-5005

Sample ID	Region	Plantation	Block	Planting Year	Planting Generation	Tree Symptomes	Tree Genetic Background
363	Sumatra	Mata Pau	MP43	1995	Fourth	medium	IL-5005
364			MP43	1995	Fourth	medium	IL-5005
366			MP50	1998	Second	important	IL-5005
367			MP50	1998	Second	light	IL-5005
368			MP50	1998	Second	medium	IL-5005
370			MP50	1998	Second	important	IL-5005
371			MP50	1998	Second	medium	IL-5005
372			MP50	1998	Second	medium	IL-5005
374		Tanah Gambus	TG10	2003	Fourth	dead	S-6805
375			TG10	2003	Fourth	dead	S-6805
376			TG10	2003	Fourth	light	S-6805
377			TG10	2003	Fourth	healthy	S-6805
378			TG10	2003	Fourth	light	S-6805
379			TG10	2003	Fourth	light	S-6805
380			TG10	2003	Fourth	medium	S-6805
381			TG10	2003	Fourth	healthy	S-6805
382			TG18	2001	Third	important	IL-6305
383			TG18	2001	Third	light	IL-6305
384			TG18	2001	Third	important	IL-6305
385			TG18	2001	Third	dead	IL-6305
386			TG18	2001	Third	healthy	IL-6305
387			TG18	2001	Third	healthy	IL-6305
388			TG18	2001	Third	important	IL-6305
389			TG29	1989	Third	important	IY-0504
390			TG29	1989	Third	medium	IY-0504
391			TG29	1989	Third	dead	IY-0504
392			TG29	1989	Third	medium	IY-0504
393			TG29	1989	Third	important	IY-0504
394			TG29	1989	Third	medium	IY-0504
395			TG29	1989	Third	important	IY-0504
396			TG25	2001	Fourth	medium	IL-6305
397			TG25	2001	Fourth	dead	IL-6305
398			TG55	1999	Second	important	IL-6305
399			TG55	1999	Second	important	IL-6305
400			TG55	1999	Second	light	IL-6305
401			TG55	1999	Second	medium	IL-6305
402			TG55	1999	Second	medium	IL-6305
403			TG55	1999	Second	light	IL-6305
404			TG55	1999	Second	dead	IL-6305
405			TG55	1999	Second	important	IL-6305
406			TG56	2000	Second	medium	IL-6305
407			TG56	2000	Second	dead	IL-6305
408			TG56	2000	Second	important	IL-6305
409			TG56	2000	Second	important	IL-6305
410			TG56	2000	Second	important	IL-6305
411			TG71	1999	Second	dead	IL-6305
412			TG71	1999	Second	healthy	IL-6305
413			TG71	1999	Second	healthy	IL-6305

Sample ID	Region	Plantation	Block	Planting Year	Planting Generation	Tree Symptomes	Tree Genetic Background
414	Sumatra	Tanah Gambus	TG71	1999	Second	dead	IL-6305
415			TG71	1999	Second	healthy	IL-6305
417			TG71	1999	Second	healthy	IL-6305
418			TG71	1999	Second	healthy	IL-6305
419			TG69	1989	Third	important	MIXED
420			TG41	1985	Third	important	MIXED
421			TG41	1985	Third	medium	MIXED
422			TG41	1985	Third	light	MIXED
423			TG41	1985	Third	dead	MIXED
424			TG41	1985	Third	light	MIXED
425			TG41	1985	Third	important	MIXED
426			TG41	1985	Third	important	MIXED
427			TG44	1996	Second	light	SY-6204
428			TG44	1996	Second	light	SY-6204
429			TG44	1996	Second	dead	SY-6204
430			TG44	1996	Second	healthy	SY-6204
431			TG44	1996	Second	dead	SY-6204
432			TG44	1996	Second	healthy	SY-6204
433			TG44	1996	Second	dead	SY-6204
AL2		Aek Loba	AL24	1998	Second	medium	Un know
AL3			AL145	2005	Second	important	Un know
AL4			AL25	1998	Second	medium	Un know
AL5			AL24	1998	Second	important	Un know
AL7			AL145	2005	Second	important	Un know
AL8			AL137	2006	Second	important	Un know
SL01		Sei Liput	SL53	1992	Second	healthy	Un know
SL02			SL76	1994	Second	light	Un know
SL03			SL62	1993	Second	light	Un know
SL04			SL76	1994	Second	healthy	Un know
SL05			SL43	1991	Second	light	Un know
SL06			SL43	1991	Second	light	Un know
SL07			SL53	1992	Second	medium	Un know
SL08			SL45	1990	Second	light	Un know
SL09			SL79	1992	Second	light	Un know
SL11			SL63	1993	Second	light	Un know
SL12			SL93	1993	Second	light	Un know
SL13			SL10	2000	Second	light	Un know
SL15			SL62	1993	Second	healthy	Un know
SL17			SL62	1993	Second	light	Un know
SL18			SL63	1993	Second	medium	Un know
SL20			SL62	1993	Second	light	Un know
SL21			SL85	1993	Second	light	Un know
SL22			SL10	2000	Second	medium	Un know
SL23			SL12	1998	Second	light	Un know
SL24			SL45	1990	Second	medium	Un know
SL26			SL63	1993	Second	light	Un know
SL27			SL76	1994	Second	light	Un know

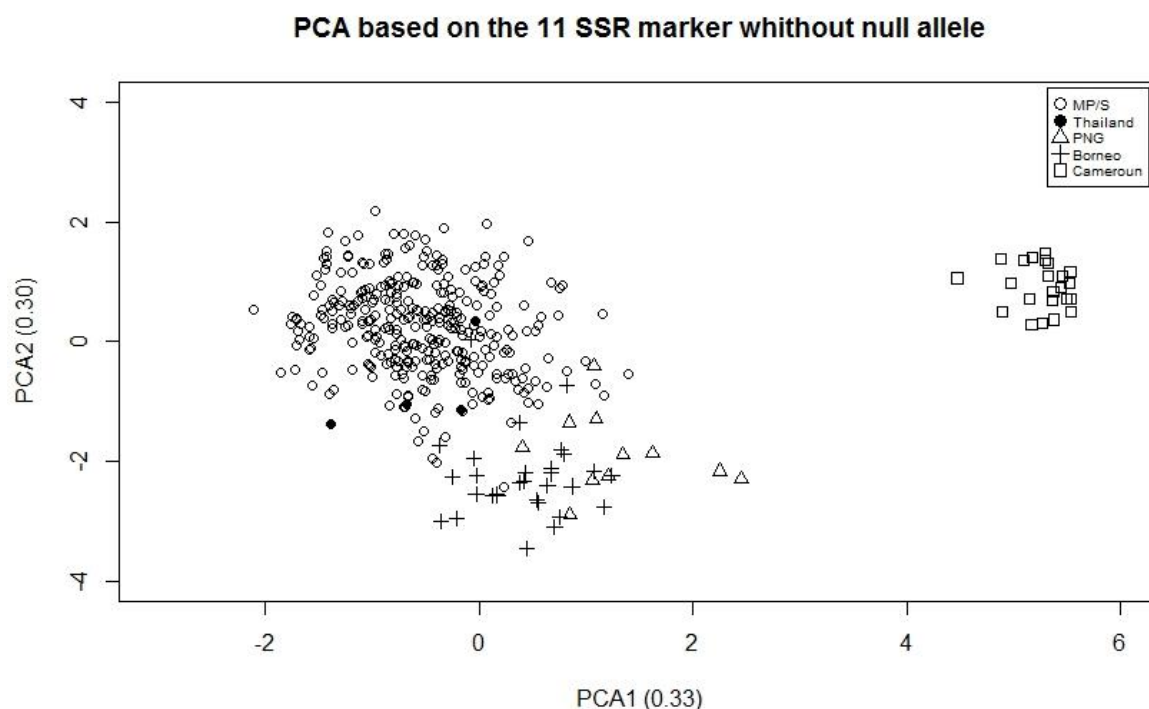
## 2. Structure mondiale des populations de *Ganoderma* en plantation de palmier à huile : résultats préliminaires



**Figure 15.** Résultats de l'ACP obtenue à partir des 11 marqueurs considérés comme sans allèles nuls pour l'ensemble des individus provenant d'Asie du Sud-Est

Au cours de l'étude précédente, les analyses portant sur les isolats provenant des missions d'échantillonnage de Malaisie péninsulaire et de Sumatra ont confirmé que ceux-ci appartiennent à une même espèce et ont montré qu'ils partagent une histoire commune. Cependant les résultats de l'article concernant le développement des marqueurs microsatellites montrent une nette différence entre les individus provenant des plantations de Bornéo et ceux du groupe MP/S.

Nous avons donc voulu savoir dans un premier temps si les échantillons envoyés par le Dr Pilloti provenant de Papouasie-Nouvelle-Guinée (PNG) partagent un profil de diversité commun à l'un des deux groupes précédemment identifiés. Ces échantillons ont été extraits et génotypés suivant le même processus que celui décrit dans Mercière et al. (2015). Les résultats obtenus ont permis de réaliser une ACP intégrant l'ensemble des individus d'Asie du sud-est (figure 15). Ces résultats montrent que les individus de Thaïlande se regroupent au sein du groupe MP/S et montrent la formation par les individus de PNG d'un groupe se rapprochant du groupe de Bornéo et mais restent différent du groupe MP/S. Malgré la formation de ces groupes indépendants on remarque que certains individus de Bornéo et de PNG ségrégent à la fois avec le groupe MP/S. Ces individus pourraient être apparentés à des migrants supposant la possibilité d'un transfert de gènes entre ces deux groupes. Ces résultats montrent une grande diversité des individus de *G. boninense* à travers l'Asie du sud-est et suggèrent que les programmes d'amélioration devraient intégrer cette diversité dans le test des individus tolérants.



**Figure 16.** Résultats de l'ACP obtenues à partir des 11 marqueurs considérés comme sans allèles nuls pour l'ensemble des individus de la collection.

Concernant l'analyse des isolats Camerounais, les résultats d'ACP obtenus pour l'ensemble des individus de notre collection ont permis de distinguer la formation d'un groupe d'individus intégrant les individus camerounais. Ce groupe apparaît comme distinct des individus provenant d'Asie du sud-est (Figure 16). Cette distinction ainsi que la mauvaise amplification de certains marqueurs microsatellites sur un grand nombre des individus camerounais (5 individus amplifiés sur le locus 6d sur un total de 32 individus et 0 pour les loci 17b et 11d), soulève la question de l'appartenance de ces individus Camerounais à l'espèce *G. boninense*. Afin de résoudre cette question, des marqueurs classiques de phylogénie (rbcL, ITS, ...) ont été séquencés sur les individus camerounais, les individus de Bornéo, les individus de PNG et un sous-échantillonnage d'individu du groupe MP/S.

Ces résultats feront l'objet d'une étude plus approfondie et d'une publication après thèse.

**CHAPITRE IV :**

**BASE TRANSCRIPTOMIQUE DE  
L'AGRESSIVITE DE *GANODERMA  
BONINENSE* (étude prélimaire)**



## 1. Introduction

Au cours du chapitre précédent, notre attention s'est portée sur la diversité génétique *a priori* neutre du pathogène. Cependant, le pathologiste est plus directement intéressé par d'autres composantes de la diversité génétique des pathogènes, et notamment sa diversité d'agressivité. Afin de bien comprendre le concept d'agressivité d'un pathogène, il est nécessaire de définir la notion de pouvoir pathogène. En phytopathologie, le pouvoir pathogène se décline traditionnellement en une composante qualitative qui correspond à la capacité d'un agent pathogène à infecter un hôte donné et à causer une maladie, et une composante quantitative correspondant au degré de dommages infligés à l'hôte par l'infection du pathogène (Van der Plank 1968). Pour la majorité des phytopathologistes européens, la composante qualitative est appelée virulence et repose sur l'interaction gène pour gène (Flor 1971). La composante quantitative est quant à elle nommée agressivité (Van der Plank 1968). Au cours de ce chapitre le terme d'agressivité sera utilisé pour qualifier la composante quantitative sous forme d'une cinétique, c'est-à-dire la quantité de symptômes induit par le pathogène à l'hôte dans un laps de temps donné.

Concernant l'agressivité de *G. boninense*, Breton et al. (2006) ont démontré qu'il existe une grande diversité entre les isolats. Ainsi des isolats distincts induisent plus ou moins de symptômes indifféremment du stade de développement de la plante (graines germées ou jeunes plants au stade 3 feuilles) au bout d'un temps donné pour tous les individus. La problématique développée dans ce chapitre concerne les facteurs génétiques responsables cette diversité d'agressivité. Jusqu'à présent, la plupart des études s'intéressant à l'interaction de l'hôte et de son pathogène, dans le pathosystème *G. boninense/Elaeis guineensis*, n'étaient développées que du côté de l'hôte. Ces études ne s'intéressaient qu'aux gènes exprimés par l'hôte au cours de l'infection (Tan et al. 2013), pouvant être reliés à de potentiels sources de résistance. Afin de bien comprendre une interaction, il est important de connaître le mode d'infection de l'agent pathogène et de déterminer les gènes qu'il mobilise au cours du processus d'infection. Dans le cas de *G. boninense*, l'interaction entre l'hôte et le pathogène est supposée non spécifique comme pour la plupart des pourridiés, c'est-à-dire qu'il n'existe pas de protéines d'avirulence spécifiques d'un gène de résistance (pas d'interaction gène pour gène). On ne peut donc cibler nos recherches sur l'identification d'une seule protéine par des moyens biochimiques. Les profils variables d'agressivité ne doivent donc pas être reliés à une variation allélique de la protéine d'avirulence mais plutôt à l'expression différentielle de gènes impliqués dans le processus de colonisation.

L'étude de cette variation d'agressivité par une approche transcriptomique va nous permettre d'explorer de façon qualitative et quantitative les gènes différentiellement exprimés entre un isolat très agressif et un isolat peu agressif. Afin d'apporter un support statistique suffisant à l'interprétation des données de transcriptomique, il est nécessaire de pallier les biais induits par la variabilité intrinsèque de chaque isolat et les biais induits par les outils utilisés. Ainsi il sera nécessaire de produire des répliques biologiques permettant la correction de la variabilité intra-isolat, ainsi que des répliques techniques permettant de corriger les variations pouvant être induites par le processus de séquençage.

Le but de ce chapitre est donc de répondre à la problématique concernant les bases transcriptomiques de l'agressivité de *G. boninense*. Cette problématique intègre deux questions majeures : existe-t-il des gènes différentiellement exprimés entre isolat agressif et non agressif ? Si oui quels sont-ils et ont-ils des fonctions connues ? Nous détaillerons le choix du matériel fongique ainsi que les méthodologies utilisées pour répondre à ces questions. Au cours de cette étude deux plans d'expérience indépendants ont été successivement mis en place due à l'échec rencontré au cours de l'exploitation des échantillons obtenus au cours de la première expérience. Cependant, le premier plan d'expérience a fait partie intégrante de mon travail de thèse et il me semble important de l'intégrer au manuscrit afin de tirer des conclusions de cet échec. Deux plans d'expérience seront donc décrits : un premier plan d'expérience développé en pré-pépinière, basé sur les tests précoces de résistance à la BSR, et un second en milieu contrôlé, stérile, développé sur le principe de culture du mycélium en milieu nutritif.

## 2. Expérimentation en préépinière

### 2.1. Matériel et Méthodes

#### 2.1.1. Choix des isolats candidats

**Tableau 1. Niveau d'agressivité et origine des 5 isolats choisis pour le premier plan d'expérimentation**

Isolats	Origine	Coordonnées GPS/Plantation	Estimation de l'agressivité des isolats inoculés en préépinière (GLM-regression logistique)		
			Moyenne (%)	Bornes à 95 % de l'intervall de confiance de la moyenne	
				Inférieur (%)	Supérieur (%)
I65	sporophore		0.4556	0.3921	2.7367
GI3	sporophore	N3 14.045 E99 25.441	4.7119	2.5867	5.4096
I5	stipe (apical)	N3 11.528 E99 20.519	12.4526	3.5903	4.7699
NJ3	sporophore	Tanah Gambus	20.1798	3.1005	3.5027
I13	sporophore	N3 10.927 E99 23.507	33.1385	5.0002	5.4121

Basé sur la mycothèque utilisée en test précoce par notre partenaire PT SocfIndo, cinq isolats possédant des profils d'agressivité contrastée, allant du très peu agressif au très agressif, ont été choisis : I65, GI3, I5, NJ3, I13. L'agressivité de ces cinq isolats ainsi que des informations sur leur provenance sont décrites dans le tableau 1. Le choix d'individus aux profils d'agressivités contrastés devait nous permettre de maximiser les chances d'observer également des profils d'expression très différents entre les individus. L'intégration de l'isolat NJ3 dans le plan d'expérimentation de l'étude transcriptomique nous permet à la fois de disposer d'un isolat à forte agressivité mais également de l'isolat de référence dont le génome a été séquencé auparavant. L'obtention des données de transcriptomique de l'isolat NJ3 permettra de réaliser l'annotation du génome de référence, ce qui en retour facilitera l'étude des données transcriptomiques de l'ensemble des individus.

#### 2.1.2. Plan d'expérience



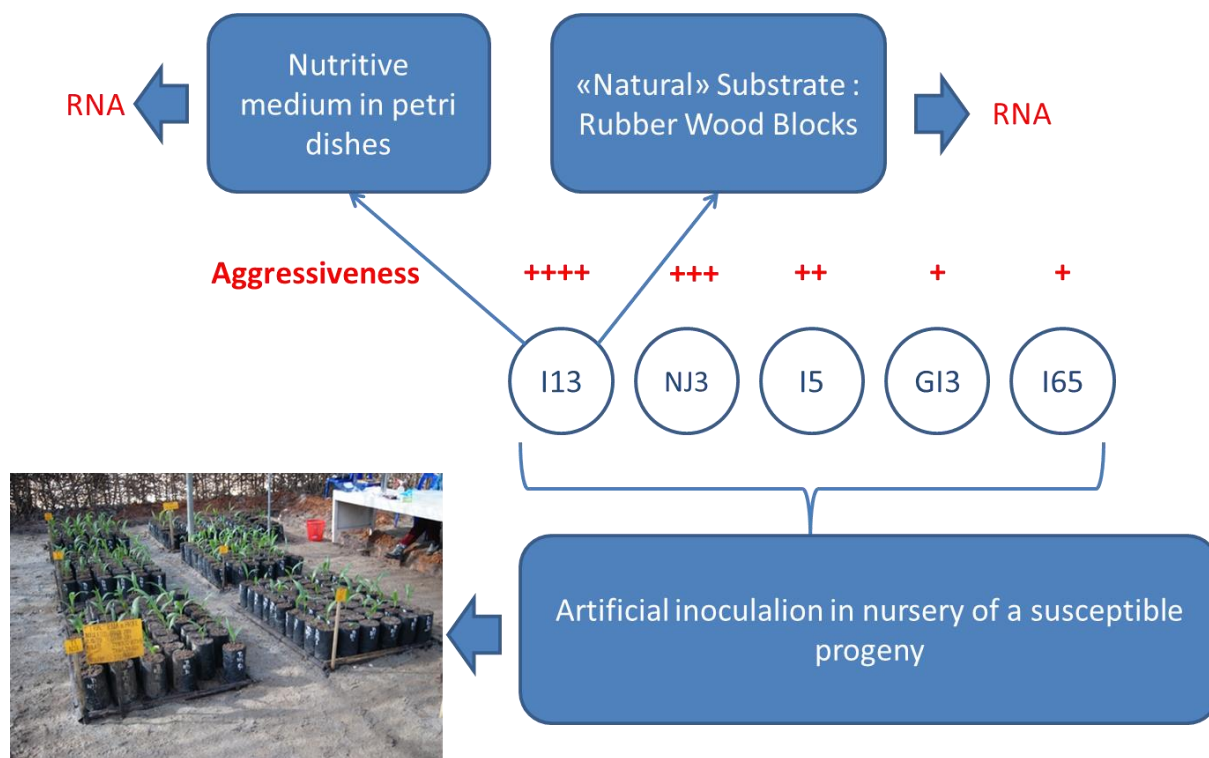
**Figure 17. Récupération du mycélium colonisant les blocs de bois d'hévéa. A) Blocs colonisés dans leur sac polyvinyle. B) Blocs colonisés avant découpe. C) extractions des tissus blanchâtre signe de colonisation par le mycélium. D) et E) conditionnement des échantillons en tube cryoconservation pour conservation dans l'azote liquide**

Le plan d'expérience imaginé repose sur trois conditions de culture pour chacun des 5 isolats.

Un premier milieu de culture stérile en boîte de Pétri, constitué de PDA (PDA Sigma, 39 g.L<sup>-1</sup>) et de chloramphénicol (0,5 g.L<sup>-1</sup>) permet l'obtention d'un mycélium pur pour lequel l'isolat n'est en contact avec aucun support ligno-cellulosique. Les boîtes de Pétri ensemencées sont ensuite placées à l'obscurité dans un incubateur à 27-28°C. Ce premier milieu a pour but d'observer si les gènes exprimés en présence d'un support ligno-cellulosique, ou de son hôte, sont exprimés de façon constitutive ou sont spécifiquement exprimés en réponse à la détection par le champignon de composés ligno-cellulosiques. Cinq répétitions par isolats sont utilisées afin de s'affranchir de la variance intra-isolat. Les échantillons de mycélium sont stockés dans l'azote liquide en attendant l'extraction. L'ARN des isolats est extrait du mycélium après une semaine de culture.

Le second milieu est constitué par des blocs de bois d'hévéa, qui sont utilisés comme supports d'inoculation lors des tests précoces. Le bois d'hévéa a été ainsi préféré au stipe de palmier pour sa rémanence supérieure dans le sol. On place des blocs d'hévéa contenus dans des sacs de polyvinyle stérile contenant 1 à 2 cm d'eau stérile, que l'onensemence à l'aide de cubes de gélose provenant de boîte de Pétri contenant l'isolat pur. Les blocs de bois ensemencés sont

ensuite placés à l'obscurité dans une salle thermo-régulée à 27-28°C. Les isolats sont donc mis en présence d'un support mimant les conditions saprophytiques qu'il pourrait développer sur des débris de bois en plantation. Ce support devait permettre de déterminer si les gènes exprimés en conditions d'infection de l'hôte sont identiques à ceux exprimés en condition saprophytique. De la même façon que pour le premier milieu, chaque isolat est cultivé sur cinq blocs d'hévéa. Au bout de trois mois, des fragments de bloc d'hévéa colonisés ont été récupérés et plongés dans l'azote liquide, afin d'en extraire l'ARN (Figure 17).



**Figure 18. Schema du premier plan d'expérimentation utilisant les cinq isolats d'agressivité variable**

Le troisième milieu est l'hôte lui-même. Des plants dotés d'un fonds génétique relativement homogène sont utilisés pour l'inoculation des cinq isolats. En l'absence de clones, les graines de palmiers sont issues d'un même croisement contrôlé afin de limiter la variabilité pouvant être imputée aux variations génétiques du support. Chaque isolat est inoculé sur un total de soixante plantes pendant une durée de quatre mois (soit environ un mois après l'apparition des premiers symptômes de BSR). Un nombre important de plantes sont utilisées afin de disposer d'au moins 3 plantes présentant des symptômes pour les 5 isolats. Les plants présentant des symptômes foliaires de BSR sont récupérés et les racines en contact avec le bloc d'hévéa sont prélevées (cf photo figure 5 du chapitre 1) en découpant la racine à partir du point de pénétration du mycélium (tissus nécrosés brun foncé) jusqu'au collet. Ces racines sont directement plongées dans l'azote liquide et cinq racines par isolat seront utilisées en vue d'une extraction d'ARN. Ces extraits d'ARN présenteront à la fois des transcrits de palmier et de *G. boninense* les transcrits obtenues après séquençage seront criblés à l'aide du génome de référence de *G. boninense*.

Le plan global de l'expérimentation est résumé dans la figure 18. La récolte des échantillons de racines infectées et de blocs d'hévéa colonisés a été effectuée au laboratoire de phytopathologie de Tanah Gambus en Indonésie.

### 2.1.3. Extraction d'ARN

J'ai utilisé le protocole d'extraction d'ARN développé par Saïdi et al. (2009) avec une adaptation : l'étape de précipitation au chlorure de sodium initiale (NaCl) à 5M a été remplacé par une précipitation au chlorure de lithium (LiCl) à 3M. Ce protocole a été testé sur les trois milieux pour l'isolat de référence NJ3 avant d'extraire les échantillons provenant des autres isolats. Les échantillons ont été ensuite contrôlés à l'aide d'un spectrophotomètre par observation des ratios 260/230 nm et 260/280 nm qui doivent être supérieur à 2 si les ARN sont purs. La concentration des ARN a été mesurée également à l'aide d'une micro-cuve au spectrophotomètre.

## 2.2. Résultats

L'adaptation du protocole de Saïdi et al. (2009) a permis d'obtenir des ARN plus purs et plus concentré que le protocole original pour l'isolat NJ3. Les isolats I65, GI3, I5 et I13 ont donc été extraits à l'aide du même protocole. L'ensemble des échantillons a ensuite été contrôlé sur Bioanalyzer Agilent 2100 afin de vérifier l'intégrité des transcrits avant leur envoi pour séquençage. Cependant, les profils de migration des ARN obtenus ont montré une très forte dégradation des ARNs des échantillons obtenus (Annexe 3) à partir des blocs d'hévéa et des racines infectées, les rendant inutilisables pour le séquençage.

## 2.3. Discussion

Cette expérimentation en pré-pépinère, devait nous permettre d'obtenir l'intégralité des gènes exprimés par les différents isolats au cours du processus d'infection dans un hôte vivant. Dans la mesure où l'ARN total des tissus de racines infectées aurait dû être extrait, l'ARN de *G. boninense* se serait retrouvé mélangé aux ARN des cellules de palmier. Les séquences d'ARN de *G. boninense* auraient dû être récupérées après séquençage par blast des reads sur le génome de référence de l'isolat NJ3.

Au final les ARN extrait se sont trouvés trop dégradés et donc non exploitables pour le séquençage. Cette dégradation peut être due au temps attendu entre la réelle pénétration du mycélium dans les racines du palmier et l'apparition des premiers symptômes. De plus la pénétration du mycélium dans les cellules de l'hôte a perturbé leur intégrité et donc provoquée la dégradation de leur contenu intégrant les ARNs contenus dans le cytoplasme.

Une autre limite à cette expérimentation est la faible abondance de cytoplasme dans les cellules de *G. boninense*, l'obtention d'une quantité d'ARN suffisante pour le séquençage aurait nécessité l'extraction d'ARN de plusieurs racines infectées due au ratio ARN fongique/ARN hôte. En complément de cette remarque, au cours du processus d'infection, le mycélium progresse le long de la racine pour atteindre la base du collet. Les cellules en divisions, donc les cellules au métabolisme le plus actif, se situent sur le front de migration du

pathogène. Le nombre de cellules en division est donc restreint et la quantité d'ARN s'en trouve restreinte également.

Pour finir, même si la terre utilisée dans les tests précoces de pré-pépinère est stérile, l'environnement dans lequel sont placées les plantules lors de l'inoculation ne l'est pas. Cette non-stérilité totale des conditions peut influencer le processus précoce de dégradation des ARNs contenus dans les échantillons prélevés.

Si l'obtention d'ARN fongique à partir de tissus racinaires infectés devait être reconduite, l'inoculation des jeunes plantules devrait être élaborée à partir d'un milieu *in vitro* (par exemple un milieu nutritif gélosé) dans un environnement clôt et stérile en laboratoire comme une chambre de culture. Ce milieu pourrait permettre l'observation de la pénétration du mycélium de *G. boninense* et ainsi permettre de prélever les tissus infectés dès les stades les plus précoces sans attendre l'apparition des premiers symptômes.

L'échec de ce plan d'expérimentation m'a donc conduit à mettre en place une expérimentation alternative en laboratoire à Montpellier.

### **3. Expérimentation en milieu nutritif stérile**

#### **3.1. Matériel et Méthodes**

##### **3.1.1. Plan d'expérience et choix des isolats**

Pour des raisons de temps et de coût, seuls trois des cinq isolats utilisés dans le plan d'expérience précédent ont été conservés : un isolat très peu agressif (GI3), l'isolat de référence NJ3 et un isolat très agressif (II3). Ces trois isolats ont été cultivés sur deux milieux dans des erlenmeyers de 250 ml : un milieu liquide Czapek (CZ) et un milieu semi-liquide Czapek appauvri en sucre et adjuvé en sciure de stipe palmier à huile (CZOP).

Le milieu CZ est composé par les éléments suivants : saccharose (5g.l<sup>-1</sup>), Tryptone (2,2065 g.l<sup>-1</sup>), K<sub>2</sub>HPO<sub>4</sub> (1 g.l<sup>-1</sup>), KCl (0,5 g.l<sup>-1</sup>), MgSO<sub>4</sub> · 7H<sub>2</sub>O (0,5 g.l<sup>-1</sup>), FeSO<sub>4</sub> · 7H<sub>2</sub>O (0,01 g.l<sup>-1</sup>) et 1ml.l<sup>-1</sup> d'une solution de microéléments composée de ZnSO<sub>4</sub> · 7H<sub>2</sub>O (1 g.l<sup>-1</sup>), de CuSO<sub>4</sub> · 5H<sub>2</sub>O (0,5 g.l<sup>-1</sup>) et de MnSO<sub>4</sub> · H<sub>2</sub>O (0,5 g.l<sup>-1</sup>). Le pH de cette fraction liquide est ajusté à 4 avec de l'HCl (5 M). Cinquante millilitres de cette solution sont réparties dans chaque erlenmeyer.

Concernant le milieu CZOP, la concentration en saccharose a été réduite à 5 g.l<sup>-1</sup>. Des copeaux de stipe de palmiers à huile préalablement séchés pendant plusieurs jours à 70°C nous sont fournis par notre partenaire Indonésien PT Socfindo. Dès réception, ces copeaux sont broyés à l'aide d'un broyeur (Retsch, ZM 200) muni d'un filtre de diamètre de 0,5 mm. Trois grammes de la sciure sèche de palmier ainsi obtenue sont alors déposés dans chaque erlenmeyer avant l'ajout des composants du milieu Czapek.

Les erlenmeyers contenant les milieux ont été stérilisés 20 minutes à 121°C puis laissés une nuit à température ambiante avant d'être ensemencés. L'ensemencement de chaque Erlenmeyer est réalisé sous hotte à flux laminaire par 3 fragments mycéliens (5 mm<sup>3</sup>) provenant d'une culture fongique pure âgée de 15 jours et cultivée sur PDA. Les Erlenmeyers inoculés sont incubés à l'obscurité dans une étuve à 28°C durant 7 jours.



### 3.1.2. Extraction d'ARN

Les premières extractions ont été réalisées à l'aide du protocole mis en place dans le plan d'expérimentation précédent. Les premiers résultats ayant révélé un taux de polysaccharides trop élevé, un nouveau protocole développé par Falcao et al. (2008) a été utilisé. Ce protocole ayant été développé pour l'extraction d'ARN d'algues rhodophytes très riches en polysaccharides (principalement utilisé pour l'extraction d'agar), il s'avère être adapté pour *G. boninense*, lui aussi très riche en polysaccharides.

### 3.1.3. Séquençage des transcrits

Les échantillons d'ARN ont été séquencés par séquençage RNAseq illumina avec un protocole pair-end de 2 x 100 nucléotides (nt) et une taille d'insert de 180 nt. Ce séquençage a été réalisé sur séquenceur HiSeq 2500. Un nombre total de 3 échantillons par condition de culture pour chaque isolat (nommé répliquats biologiques) ont été séquencés, ce qui représente un total de 18 échantillons d'ARN séquencés (3 réplicats biologiques x 2 milieux de culture x 3 isolats). Chaque échantillon a été séparé en deux pools d'ARN ayant été séquencés de façon indépendante (réplica technique). Tous ces échantillons ont été répartis aléatoirement sur 4 lignes de séquençage (run) de sorte que les pools préparés pour les réplicats techniques ne se retrouvent jamais sur la même ligne de séquençage. Ce procédé permet de s'affranchir de la variabilité liée aux lignes de séquençage et permet de normaliser cet effet sur l'ensemble des échantillons séquencés.

### 3.1.4. Assemblage et analyse des données RNAseq

L'assemblage a été réalisé grâce au logiciel Burrows-Wheeler Aligner (BWA) (Li et al. 2009) en utilisant les paramètres par défaut. Les transcrits obtenus après assemblage ont été annotés avec le logiciel BLAST (Altschul et al. 1990). La couverture de chaque transcrit a ensuite été obtenue par comptage du nombre de reads s'alignant sur chaque transcrit pour chacun des 18 échantillons (18 fichiers de comptage). Les travaux d'assemblage, d'annotation et de comptage ont été réalisés par nos collaborateurs de l'équipe GENOTOUL de l'INRA de Toulouse. J'ai ensuite procédé aux analyses de ces résultats de ce comptage grâce au package EdgeR (Robinson et al. 2010) du logiciel R. Après une étape de normalisation des comptages pour chacun des 18 échantillons 2 types de comparaison ont été ciblé afin de répondre aux questions de recherche : une première comparaison basée sur le nombre de transcrits différentiellement exprimés pour un isolat entre les deux conditions de culture et une deuxième comparaison entre isolats pour une même condition de culture. Les significativités des gènes obtenues comme différentiellement exprimés sont basées sur un test non paramétrique de FDR (False Discovery Rate) fixé à une valeur de 5% (0.05).

### 3.2. Résultats et Discussion

#### 3.2.1. Obtention d'ARN pur

**Tableau 2. Contrôle qualité et quantité des échantillons séquencés**

	rin	concentration (ng.µL <sup>-1</sup> )
GI3 cz1	8,7	291
GI3 cz3	9,4	1512
GI3 cz4	9,9	2080
GI3 czop1	7,6	296
GI3 czop2	7,3	169,7
GI3 czop3	6,8	648
I13 cz1	8,7	141
I13 cz3	8,4	347
I13 cz4	8,5	630
I13 czop1	5,5	22
I13 czop2a	6,8	167
I13 czop3	6	30
NJ3 cz1	6,8	526
NJ3 cz2	6,6	371
NJ3 cz4	6,8	511
NJ3 czop2a	8,1	1180
NJ3 czop3a	8,3	1923
NJ3 czop4a	7,8	1353

Les ARN obtenus à l'aide du protocole développé par Falcao et al. (2008) se sont révélés de très bonne qualité. Les profils de migration obtenus sur Bioanalyzer à partir de chaque échantillon n'ont révélé aucune dégradation des ARN et aucune contamination par des composés polluants (comme les polysaccharides). En effet, la présence de composés du type polysaccharidique empêche une bonne migration des ARNs sur puce Bioanalyzer en bouchant les microcanaux. Les concentrations obtenues se situent entre 22 ng.µL<sup>-1</sup> et 2080 ng.µL<sup>-1</sup> (Tableau 2). L'indice de qualité d'ARN délivré par le Bioanalyzer et appelé RIN (RNA Integrity Number) reflète l'intégrité de l'échantillon testé. Les valeurs obtenues pour les 18 échantillons se situent entre 5.5 et 9.9. Ces résultats montrent des ARNs de bonne qualité, non dégradés et une quantité de matériel suffisante pour un séquençage de type HiSeq Illumina. Des exemples de profils de migration bioanalyzer sont présentés dans l'annexe 4. Les nouvelles techniques de séquençage étant très sensibles à la qualité des échantillons, il est primordial de fournir des ARNs de bonne qualité et en quantité suffisante afin d'éviter une étape de préséquencage consistant à amplifier les transcrits par PCR. Cette étape pourrait introduire un biais quant aux analyses de différence d'expression entre deux échantillons.

#### 3.2.2. Première analyses des résultats de séquençage

Le séquençage Hiseq Illumina est une technique de séquençage pouvant générer un total de 350 à 500 millions de reads par run et par ligne de séquençage pour une longueur de reads

**Tableau 3. Résumé des fichiers de comptage pour chaque échantillon séquencé**

files	group	description	Medium	lib.size	norm.factors
NJ3czop4a.txt	NJ3	NJ3czop4a	czop	76497459	1
NJ3czop3a.txt	NJ3	NJ3czop3a	czop	77873025	1
NJ3czop2a.txt	NJ3	NJ3czop2a	czop	64496459	1
NJ3CZ4.txt	NJ3	NJ3CZ4	cz	77881251	1
NJ3CZ2.txt	NJ3	NJ3CZ2	cz	96575902	1
NJ3CZ1.txt	NJ3	NJ3CZ1	cz	70493901	1
I13czop3.txt	I13	I13czop3	czop	83297977	1
I13czop2a.txt	I13	I13czop2a	czop	70201594	1
I13czop1.txt	I13	I13czop1	czop	48845444	1
I13cz4.txt	I13	I13cz4	cz	99470764	1
GI3cz3.txt	GI3	GI3cz3	cz	61653355	1
GI3cz2.txt	GI3	GI3cz2	cz	81249758	1
GI3cz1.txt	GI3	GI3cz1	cz	80362653	1
I13cz3.txt	I13	I13cz3	cz	65595612	1
I13cz1.txt	I13	I13cz1	cz	73436419	1
GI3czop3.txt	GI3	GI3czop3	czop	85268134	1
GI3czop1.txt	GI3	GI3czop1	czop	110492300	1
GI3czop2.txt	GI3	GI3czop2	czop	86100552	1

pouvant atteindre 250 nt maximum (<http://www.biotech.illinois.edu/>). Le nombre moyen de reads obtenus pour les 18 échantillons est de 78 millions de reads avec un maximum de 110 millions de reads obtenus pour GI3czop1 et un minimum de 49 millions de reads pour I13czop1 (Tableau 3). Ces résultats sont donc de bonne qualité considérant la quantité de reads obtenue (350 millions de reads répartis sur 4 à 5 échantillons). L'ensemble de ces reads a permis d'obtenir l'assemblage de 59 198 contigs correspondant à 59 198 transcrits différents dont 43 195 n'ont pas été identifiés par blast. Ce nombre de transcrits est plus élevé que les 12 646 transcrits assemblés par Chen et al. (2012) pour *G. lucidum*. La comparaison de ces transcrits avec les bases de données NCBI par blast, nous a montré que l'espèce la plus citée pour l'annotation de nos transcrits était pour le plus grand nombre *Postia placenta*. Cette espèce correspond à un basidiomycète phytopathogène classé parmi les pourritures brunes et dont les hôtes potentiels sont des espèces pérennes forestières comme par exemple le sapin de Douglas.

Afin de proposer une interprétation biologique concernant les transcrits différentiellement exprimés, tous les transcrits sans annotation ont été supprimés des fichiers de comptage, ce qui a recentré l'analyse sur les 16003 transcrits annotés, plus cohérent avec le nombre de transcrits de *G. lucidum* (Chen et al. 2012).

La comparaison des transcrits obtenues pour I13 avec ceux de GI3 et pour les milieux CZ et CZOP a permis de mettre en évidence respectivement 2 363 et 1 482 gènes différentiellement exprimés. La comparaison entre NJ3 et GI3 pour les milieux CZ et CZOP a mis en évidence 2 402 et 2 292 gènes différentiellement exprimés. Dans le but de répondre à la première question des gènes différentiellement exprimés entre les isolats agressifs et non agressifs, l'intersection des deux milieux pour chaque isolat a été extraite ne conservant que

1 148 gènes différentiellement exprimés et commun aux deux milieux pour la comparaison entre I13 et GI3. Le même procédé utilisé pour la comparaison entre NJ3 et GI3 a permis de réduire le nombre de gènes différentiellement exprimés et commun aux deux milieux de culture à 1 485. Pour finir l'intersection entre ces deux listes a été extraite et a permis de mettre en évidence un total de 453 gènes différentiellement exprimés commun aux deux isolats agressifs comparés à l'isolat non agressif GI3. L'élimination des gènes correspondant à des protéines hypothétiques a réduit le nombre de gènes à 124. L'observation des 10 gènes les plus différentiellement exprimés pour les comparaisons I13 vs GI3 et NJ3 vs GI3 pour les deux milieux (=4 comparaisons différentes) a permis de faire ressortir 1 gène commun à ces quatre comparaisons. Ce gène correspond à une polyporopepsin proche de l'espèce *Irpex lacteus* un polypore faisant partie du groupe des pourritures blanches. La fonction de cette protéine correspond à activité de protéase aspartique utilisant les résidus aspartate pour catalyser les réactions. Ce gène pourrait être un candidat intéressant pour discriminer les isolats agressifs et non agressifs mais nécessitera des analyses complémentaires pour confirmer son implication dans l'agressivité des isolats.

### 3.3. Conclusion

Jusqu'à présent, l'étude de l'interaction hôte-pathogène et plus particulièrement des profils d'expression a été centrée sur les gènes exprimés par le palmier à huile lors de cette interaction (Tan et al. 2013, Tee et al. 2013). Les ARNs obtenus lors de cette étude ont permis le séquençage et l'annotation d'un transcriptome de 59 198 transcrits exprimés sur deux milieux de culture et par trois isolats d'agressivités différentes. Les gènes identifiés comme différentiellement exprimés entre isolats agressifs et non agressifs sont nombreux et il sera nécessaire d'effectuer une investigation plus approfondie des données afin de révéler de potentiels gènes impliqués dans la variation des profils d'agressivité de *G. boninense*.

**CHAPITRE V :**

**DISCUSSION GENERALE ET**

**PERSPECTIVES**

Le travail effectué au cours de cette thèse avait pour objectif de mieux connaître la biologie de *Ganoderma boninense*, agent responsable de la pourriture basale du stipe chez le palmier à huile. Après avoir mis au point un jeu de 17 marqueurs SSR sur la base d'une version préliminaire de la séquence du génome de *G. boninense*, je me suis intéressé tout particulièrement à sa diversité et sa structure génétique dans la zone historique de développement de la culture industrielle du palmier à huile. Les données recueillies permettent de mieux comprendre le mode de reproduction et les capacités de dispersion de *G. boninense*. Parallèlement, j'ai également cherché à mieux comprendre les bases moléculaires de la différence d'agressivité observée au champ entre isolat. Pour cela, j'ai mis en place une expérience permettant de comparer de manière quantitative les transcriptomes produits par des isolats d'agressivité contrastée qui ont été placés dans différents milieux. Cette étude, dont je n'ai présenté que des résultats préliminaires, doit permettre une meilleure compréhension de la stratégie déployée par *G. boninense* lors du processus d'infection.

## **1. Développement de marqueurs microsatellites et caractérisation de trois populations distinctes en Asie du Sud-Est**

La réduction des coûts de séquençage ces dix dernières années, a permis le développement de marqueurs microsatellites à partir de données de séquençage NGS pour des espèces modèles ou non (Cai et al. 2013, Schoebel et al. 2013 a, Schoebel et al. 2013 b). Le séquençage du génome de l'isolat NJ3 nous a permis de faire l'inventaire des séquences microsatellites contenues au sein du génome de *G. boninense* et d'estimer la taille de son génome. Au total 2 487 marqueurs ont été identifiés pour un génome estimé à environ 65 Mb soit 903 marqueurs de plus que le génome de *G. lucidum* (Qian et al. 2013) et une taille de génome supérieur d'environ 20 Mb à celui de *G. lucidum* (Chen et al. 2012). Ce résultat correspond à une différence d'environ 35 % à la fois pour la taille de génome et pour le nombre microsatellite. Ce pourcentage identique pour les deux valeurs renforce le crédit de ces résultats et surtout une évolution entre *G. boninense* et *G. lucidum* pouvant être expliquée par l'observation de nombreuses séquences dupliquées lors de l'assemblage du génome de *G. boninense*. Contrairement à la constitution de banques enrichies en séquence répétée qui ne permettent la découverte que d'un faible nombre de marqueurs microsatellites (Schoebel et al. 2013 b), l'utilisation de données de séquençage permet l'identification d'un nombre important de marqueurs microsatellites à moindre coût et dans un laps de temps plus court. Cependant, tous les marqueurs microsatellites ne peuvent être testés *in vitro* car cela nécessiterait un temps considérable malgré la disponibilité d'appareils de séquençage haut débit comme les séquenceurs capillaires. Le développement de processus de PCR *in silico* permet donc une première sélection des couples d'amorces utilisables.

Au final 17 marqueurs microsatellites ont été développés dont 11 d'entre eux semblent réellement exempts d'allèle nul et ont servi au génotypage de l'intégralité de notre échantillonnage d'Asie du sud-est, mais également aux échantillons Africains. Le génotypage de l'intégralité des échantillons nous a permis dans un premier temps de répondre à la question du nombre de groupes génétiques existant en Asie du sud-est, et enrichit notre compréhension des mécanismes de dispersion du pathogène au sein des plantations industrielle de palmier à huile.

### 1.1. Une dispersion longue distance de *G. boninense*

Nous avons évalué dans un premier temps l'informativité des marqueurs microsatellites développés en géotypant un sous-échantillonnage issu des missions de collecte réalisées en Malaisie péninsulaire et à Sumatra. Ces données nous ont révélé la présence de deux populations distinctes au sein de cet échantillonnage : une première population englobant l'intégralité des individus provenant de Malaisie péninsulaire et de Sumatra (MP/S), et une seconde regroupant les individus provenant des plantations de Bornéo. La population regroupant les individus MP/S correspond à la population pour laquelle nos efforts d'échantillonnage ont été volontairement augmentés. Cette zone possède un échantillonnage dans lequel s'emboîtent plusieurs échelles géographiques : la région (Malaisie péninsulaire, Sumatra), la plantation et le bloc. Cet emboîtement favorise l'étude de la capacité de dispersion de *G. boninense*. De plus, cette zone correspond à la zone historique de développement à grande échelle des plantations industrielles de palmier à huile en Asie du sud-est. L'étude des données obtenues pour cette population au travers d'AMOVA montrent que la population de MP/S ne présente pas de structure à une échelle géographique plus limitée (bloc ou plantation). Cette population MP/S apparaît comme plus large que les populations observées lors d'études antérieures et donne une nouvelle dimension aux résultats obtenus par Miller et al. (1999) et plus tardivement par Pilotti et al. (2003) n'ayant étudié la diversité de *G. boninense* qu'au sein d'une ou deux plantations d'une même région. L'information concernant le fond génétique du palmier sur lequel l'échantillon a été prélevé, a également permis de constater qu'il n'y avait pas de corrélation entre un géotype particulier de *G. boninense* et le fonds génétique des palmiers. Cette observation abonderait dans le sens d'une interaction non spécifique entre le palmier à huile et *G. boninense*. De plus l'observation de l'effet du nombre de générations sur la diversité des individus de *G. boninense*, ne montre aucune variation significative de la richesse allélique en fonction des individus regroupés par nombre de générations de leur plantation d'origine. Ce résultat peut indiquer soit qu'il n'y a pas d'enrichissement de la diversité génétique de *G. boninense* en plantations au fil des générations, ce qui aurait été permis par une banque d'isolats s'accumulant dans le sol, soit que la diversité en plantation est trop grande pour pouvoir détecter ce phénomène. La population MP/S nous a également permis dans un second temps de mieux comprendre le mécanisme et la capacité de dispersion de *G. boninense*.

L'observation d'une forte diversité génétique au sein du groupe MP/S, contrairement à l'étude d'autres basidiomycètes phytopathogènes comme *Armillaria mellea* (Prospero et al. 2008, Travadon et al. 2012 b), nous oriente vers une prédominance de la dispersion par spores issues de la reproduction sexuée. D'autres études comme chez *heterobasidion annosum* (Woodward et al. 1998) suggèrent une dispersion par le vent des basidiospores permettant les flux génétiques sur de très grandes distances. De tels profils de dispersion ont déjà été observés chez plusieurs autres basidiomycètes (Hallenberg et al. 2001). De plus, le faible nombre de clones identifiés et le fait qu'ils n'aient été retrouvés que sur des palmiers voisins (9 m entre chaque palmier) abondent dans le même sens. La mise en œuvre systématique de pratiques culturales visant à réduire la transmission racinaire de *G. boninense* pourrait également expliquer ce très faible nombre de clones retrouvés en plantation. Concernant la



capacité de dispersion de *G. boninense*, l'absence structure génétique et d'isolation par la distance au sein des individus provenant de Malaisie péninsulaire et de Sumatra montrent la capacité des basidiospores à franchir des bras de mer comme ici le détroit de Malacca.

Dans un troisième temps, nous avons cherché à comprendre la structure génétique des individus de *G. boninense* à une échelle plus large englobant l'intégralité des échantillons d'Asie du sud-est présents au sein de notre collection. Les résultats des analyses de structuration ont mis en évidence le regroupement des 4 échantillons provenant de Thaïlande avec le groupe MP/S, ainsi que le regroupement des échantillons de PNG au sein du groupe de Bornéo. Dans la mesure où jusqu'à présent seul des individus récoltés à l'échelle d'une plantation avaient fait l'objet d'étude de diversité, la structuration de la diversité des populations de *G. boninense* à l'échelle de l'Asie du sud-est n'avait pu être mise en évidence. Au cours des différentes analyses de structuration génétique des individus de *G. boninense*, plusieurs individus semblaient partager des caractéristiques génétiques communes à au moins deux des trois populations identifiées. Des analyses plus approfondies devront être mise en place afin de déterminer si ces individus pourraient être des migrants ou si leur ségrégation au sein d'un groupe distant résulte d'artéfacts. Le regroupement des échantillons Thaïlandais au sein du groupe MP/S devra être confirmé par l'intégration d'un plus grand nombre d'échantillons Thaïlandais au sein des analyses.

Pour finir, les études de structuration intégrant les individus provenant d'Afrique montrent qu'ils forment un groupe indépendant des individus d'Asie du sud-est. Dans la mesure où plusieurs échantillons n'ont pu être intégrés aux analyses du fait d'une non-amplification de plusieurs marqueurs microsatellite, ces échantillons Africains doivent soit avoir subi des modifications au niveau des sites complémentaires des amorces microsattellites, soit appartenir à des espèces différentes. Ce résultat nous permet donc de nous questionner quant aux espèces de *Ganoderma* présente dans les plantations africaines. Au cours de ces derniers mois de thèse des marqueurs de type barcoding (ITS, RbcL, mtSSU, ...) ont été séquencés sur les individus de notre collection et l'analyse de ces données pourra permettre de statuer sur la classification des individus récoltés.

## 1.2. Une expansion démographique ancienne

L'étude des données de génotypage pour les individus de la population MP/S a permis de retracer l'histoire démographique de *G. boninense* au sein de la zone historique de culture industrielle du palmier à huile en Asie du sud-est. Cette zone est d'une importance capitale pour la compréhension de l'histoire démographique de *G. boninense* car elle représente la zone le contact entre l'hôte et le pathogène. En effet, plus les contacts entre un pathogène et son hôte sont anciens plus l'impact de l'hôte sur la distribution des génotypes et les modifications du génome du pathogène sont importantes et donc plus facilement détectables. La comparaison entre différents scénarios retraçant les potentielles variations de taille de population ayant eu lieu depuis le contact entre l'hôte et le pathogène a permis de dégager le scénario le plus probable. Ce scénario correspond à une expansion de la population de *G. boninense* dans la zone MP/S. Néanmoins l'estimation du début de ce phénomène d'expansion ne correspond pas à la période du début du contact entre l'hôte et le pathogène

comme on aurait pu le supposer. D'après cette estimation, la période du début de l'expansion correspondrait à la fin du maximum glaciaire de la dernière ère glaciaire, il y a environ 17 000 ans. Comme beaucoup d'espèces à travers le monde, cette période a eu une forte influence sur l'évolution de la distribution des génotypes et l'évolution des génomes (Hewitt 2000). Ainsi en Asie du sud-est les forêts équatoriales ont vu leur territoire augmenter à cette période (Hewitt 2000, Raes et al. 2014). Néanmoins, les espèces au sein de ces territoires n'ont pas toutes subi une expansion de leurs populations, certaines comme les Dipterocarpaceae ont même subi une réduction de leur taille comparée à leur taille actuelle (Raes et al. 2014). *Ganoderma boninense* possédant un large spectre d'hôte a donc pu subir la réduction de taille de population de ses hôtes et donc lui aussi partir d'une taille de population réduite à cette période pour obtenir après expansion sa taille actuelle.

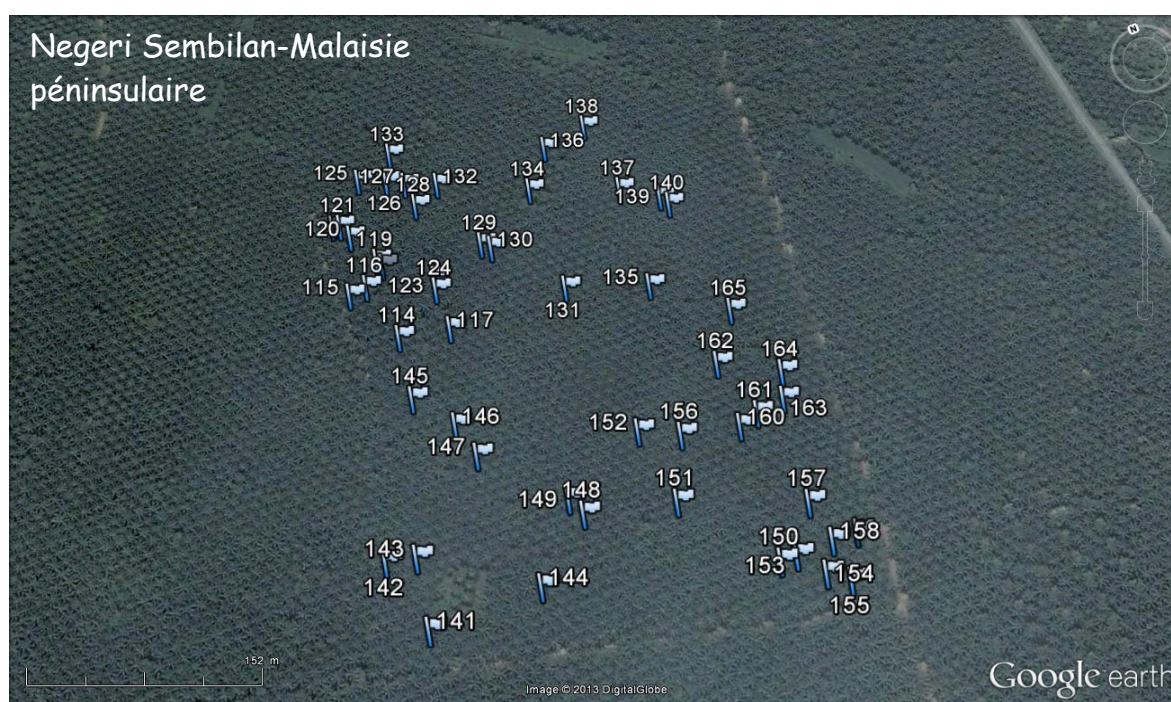
L'apparition de *Ganoderma boninense* sur palmier à huile n'a donc, au regard de ces résultats, pas laissé d'impact sur la répartition et la diversité des génotypes. L'impossibilité de détection d'un tel phénomène peut venir soit du fait qu'il n'y a réellement eu aucun impact sur l'évolution démographique de *G. boninense*, mise à part la mise à disposition d'un nouvel hôte accroissant ainsi le phénomène d'expansion, soit que la détection d'un tel phénomène de saut d'hôte nécessite l'utilisation d'un plus grand nombre de marqueurs répartis uniformément sur le génome dans des zones à évolution rapide et à évolution lente. L'utilisation d'un plus grand nombre de marqueurs microsatellites ou l'utilisation d'un très grand nombre de marqueurs SNP pourrait constituer une approche envisageable pour une exploration plus fine de l'histoire démographique de *G. boninense*. Pour finir, cet événement de saut d'hôte peut également être trop récent pour pouvoir être détecté.

## **2. Etude préliminaire de gènes différentiellement exprimés**

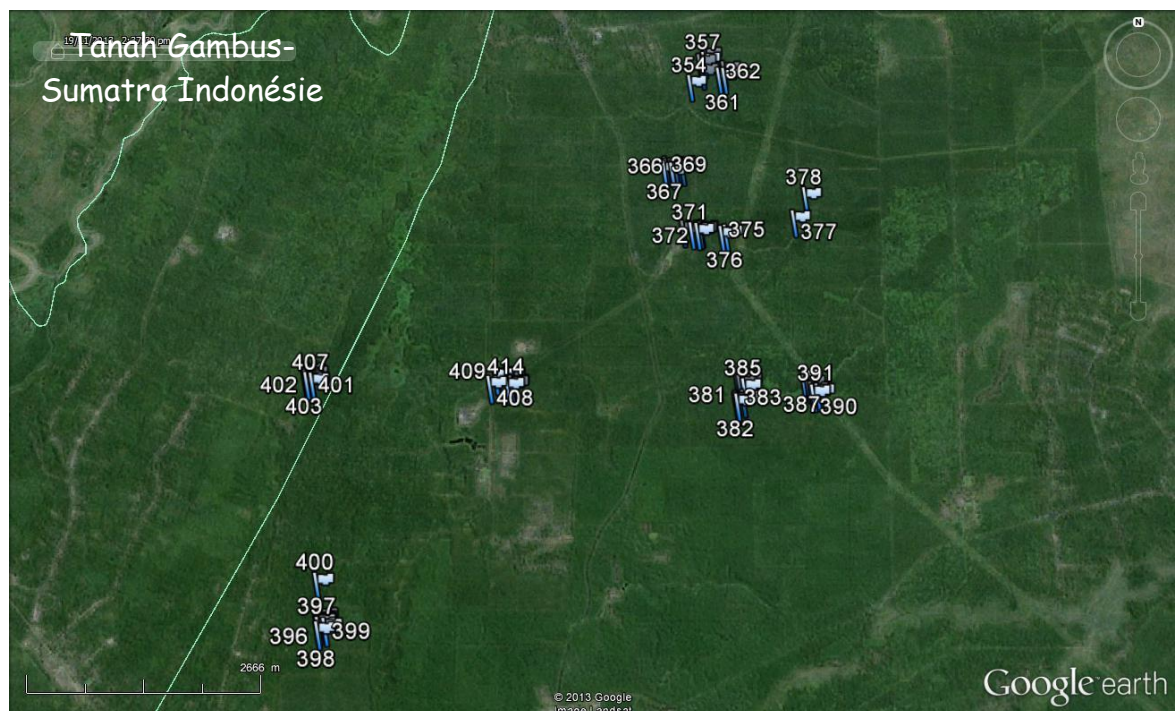
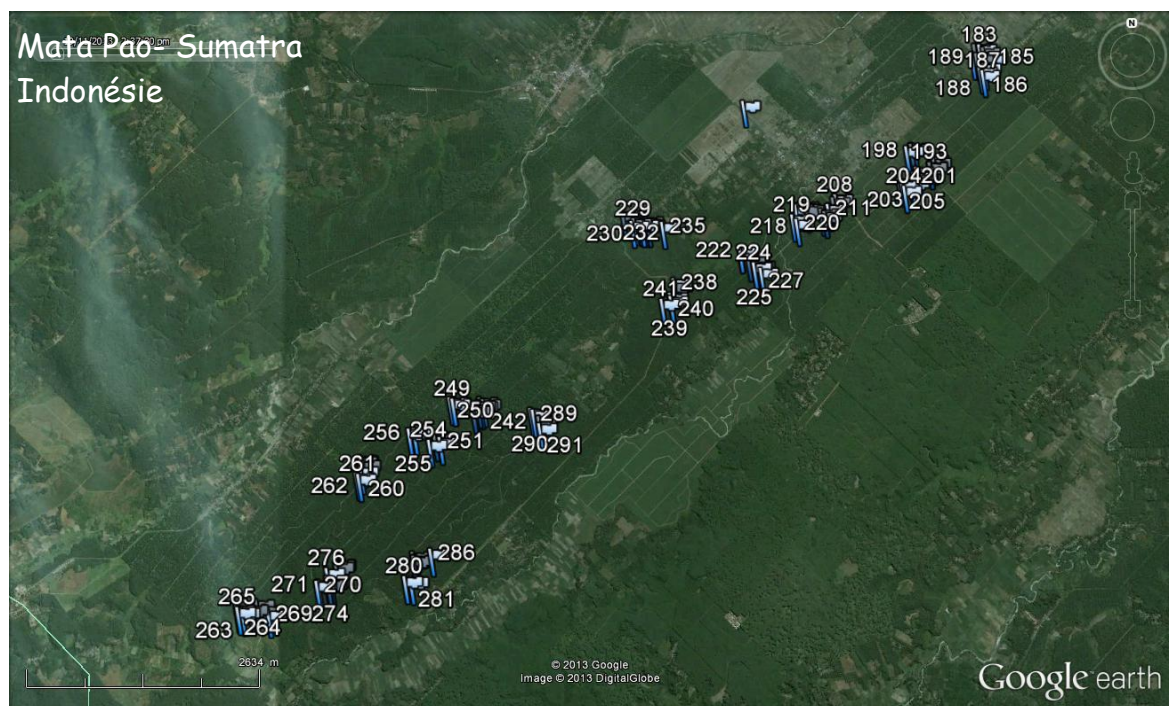
Jusqu'à présent, aucune étude ne traite de l'expression des gènes de *G. boninense* au cours de son interaction avec l'hôte ou dans sa phase saprophytique. L'étude d'isolats possédant des profils d'agressivité variable a permis l'observation de données préliminaires concernant la présence de gènes différentiellement exprimés. Ces résultats nous ont montrés que 124 gènes possédant une fonction connue sont différentiellement exprimés entre les isolats agressifs et non agressifs. L'ajout de sciure de palmier ou non dans le milieu ne semble pas influencer la variation d'expression de ces gènes. Malheureusement il est impossible de statuer si ces gènes sont exprimés au même niveau au cours de l'infection de l'hôte par le pathogène car la sciure de palmier à huile est stérilisée à l'autoclave (aucune cellule vivante) et aucun autre milieu contenant un autre type de sciure n'a été testé. Ces gènes peuvent donc être uniquement spécifiques des conditions saprophytiques. Afin de confirmer ces résultats, des analyses plus approfondies des profils d'expression devront être réalisées, et un nouveau milieu contenant un autre type de sciure (par exemple de la sciure d'hévéa) pourrait être testé ou dans un test *in vivo* développé en chambre de culture stérile.

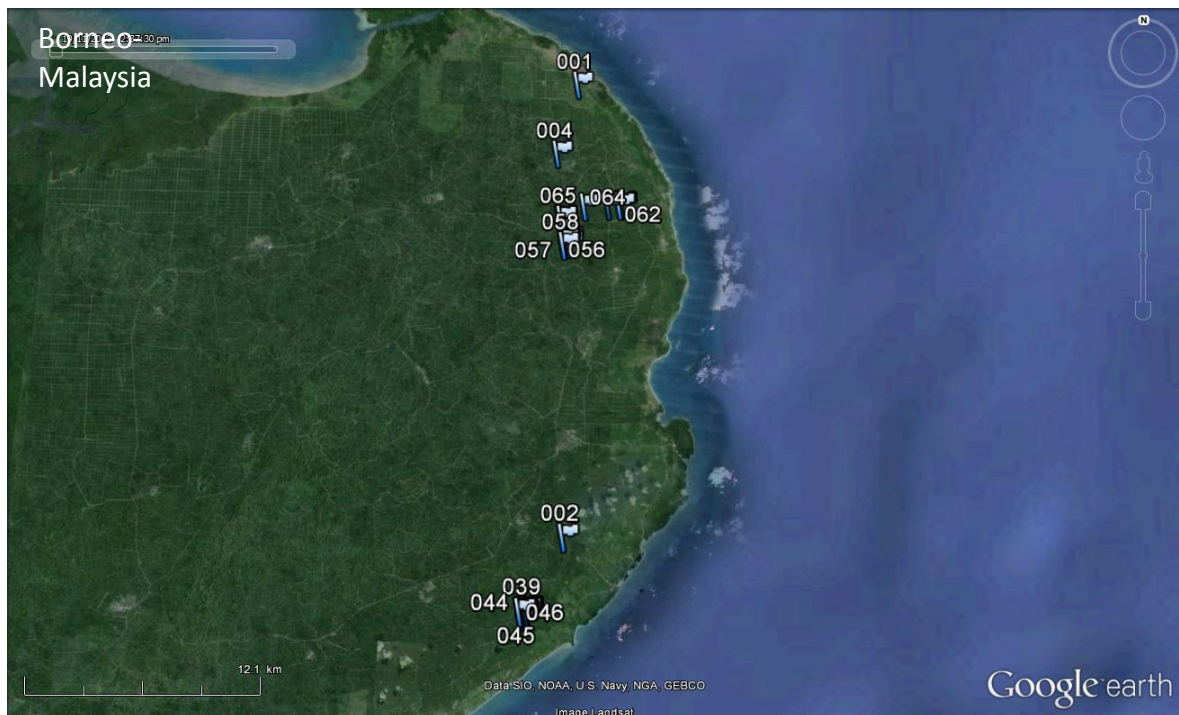
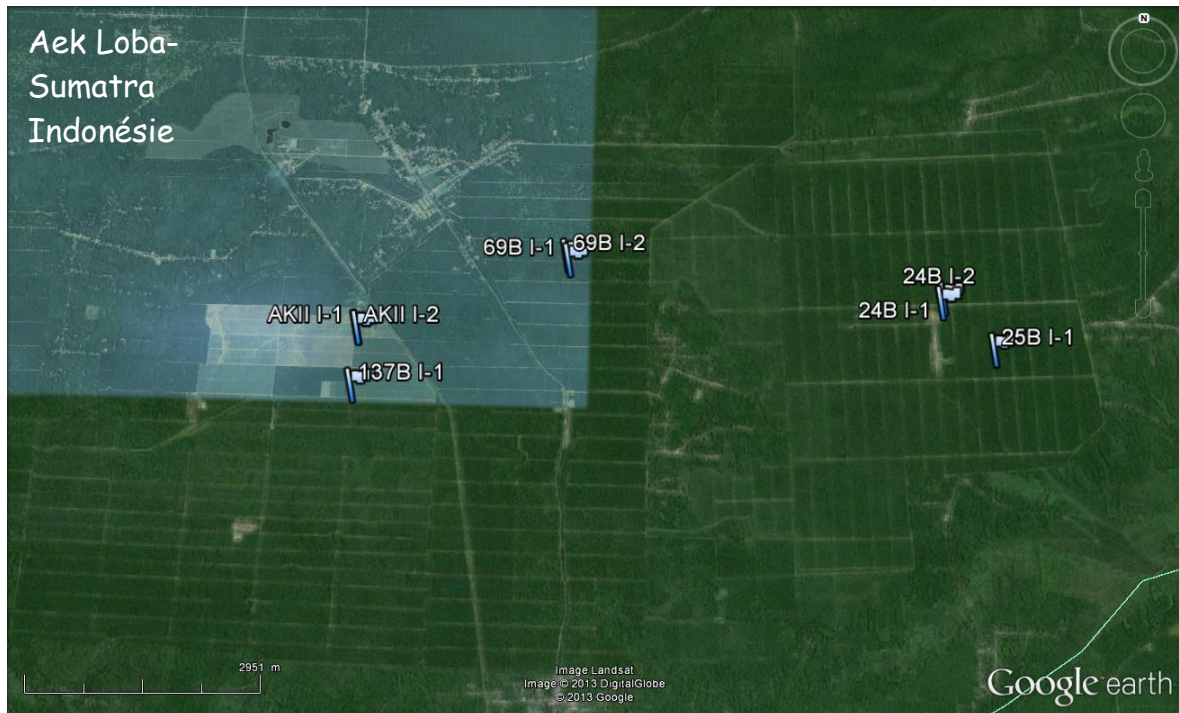
# **ANNEXES**

# **ANNEXE 1 : Carte représentant les différentes plantations échantillonnées en Asie du Sud-Est**







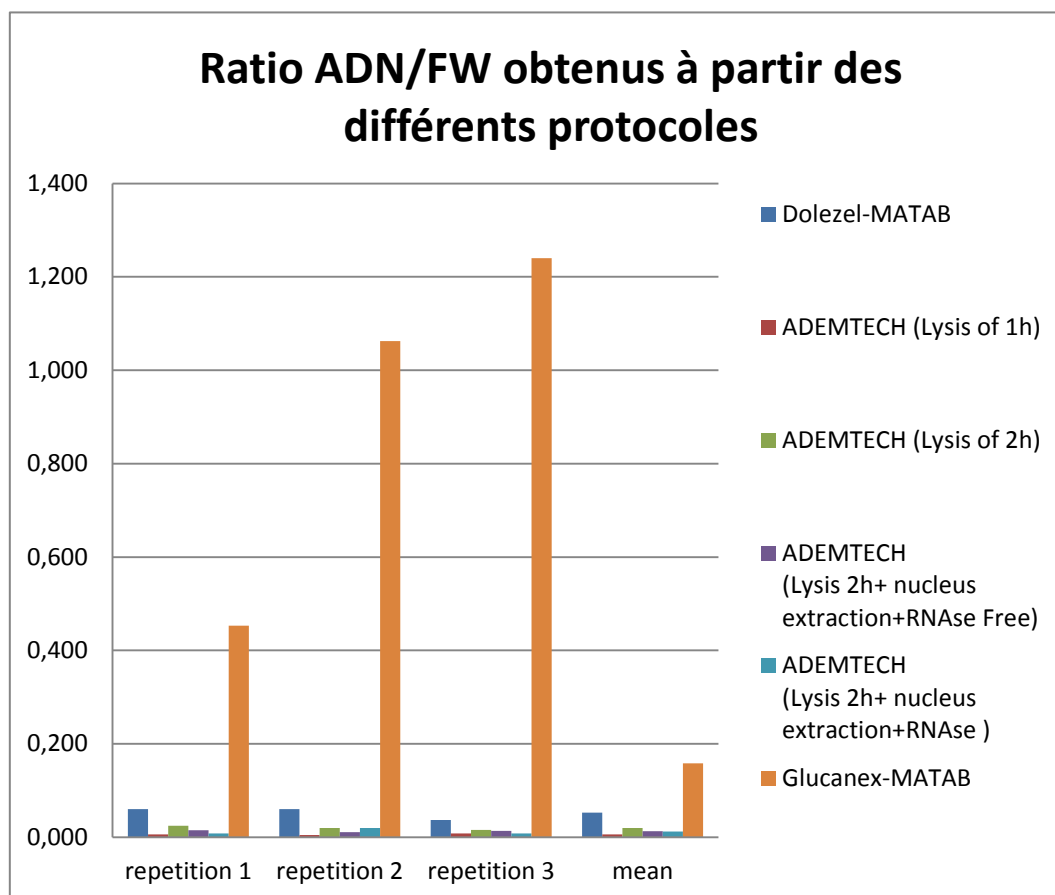
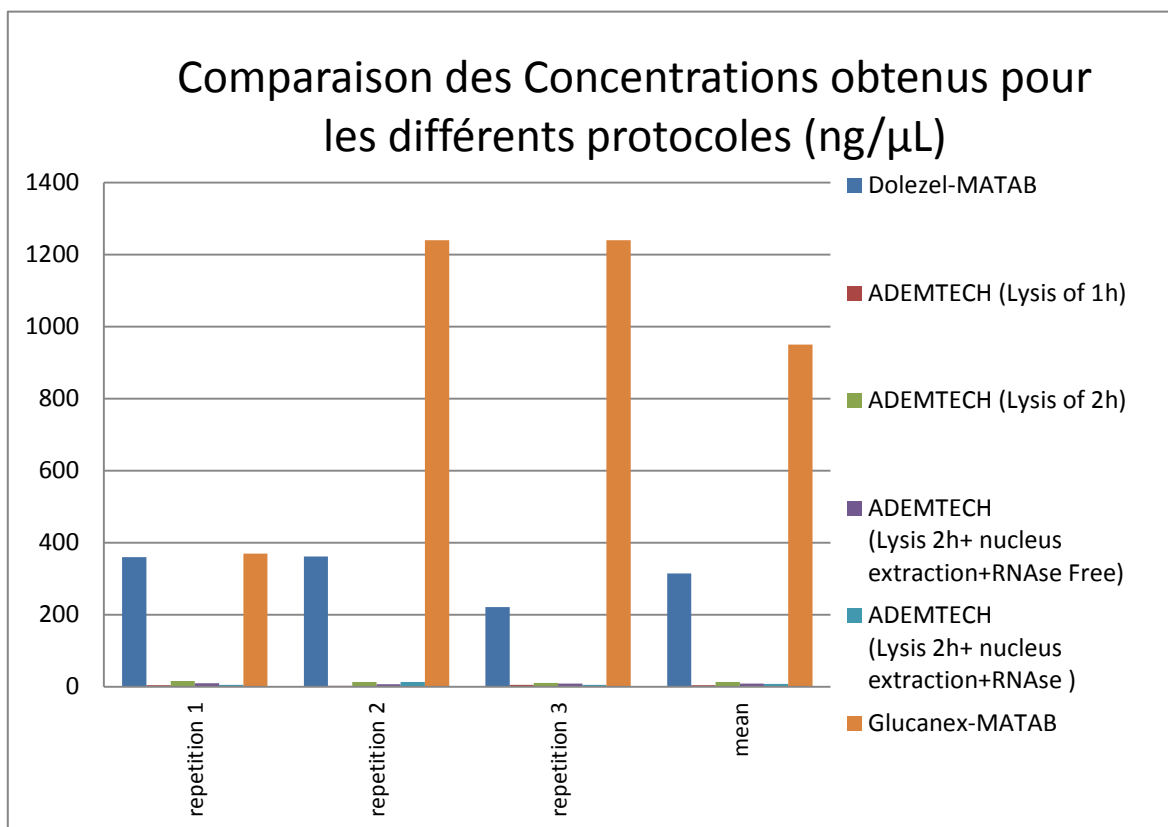




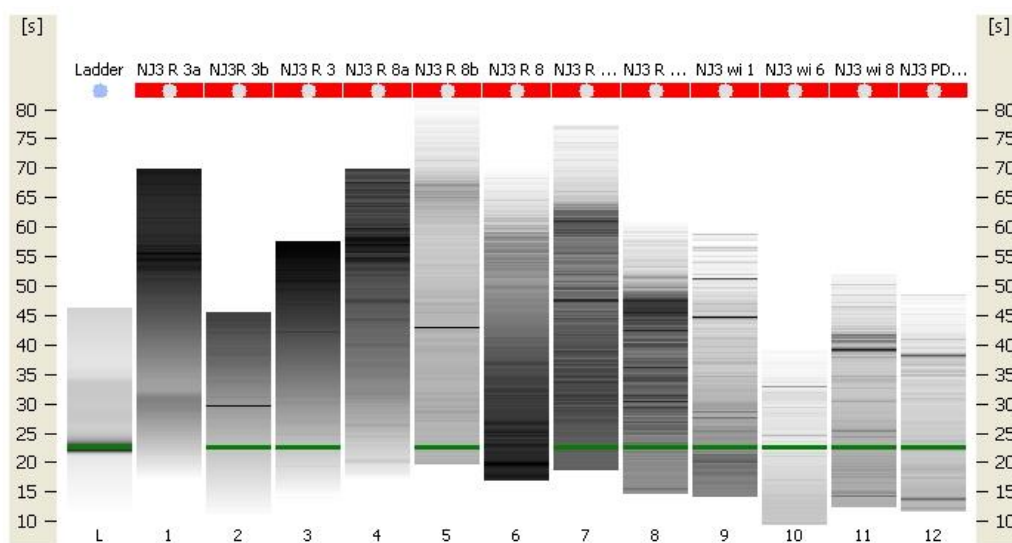
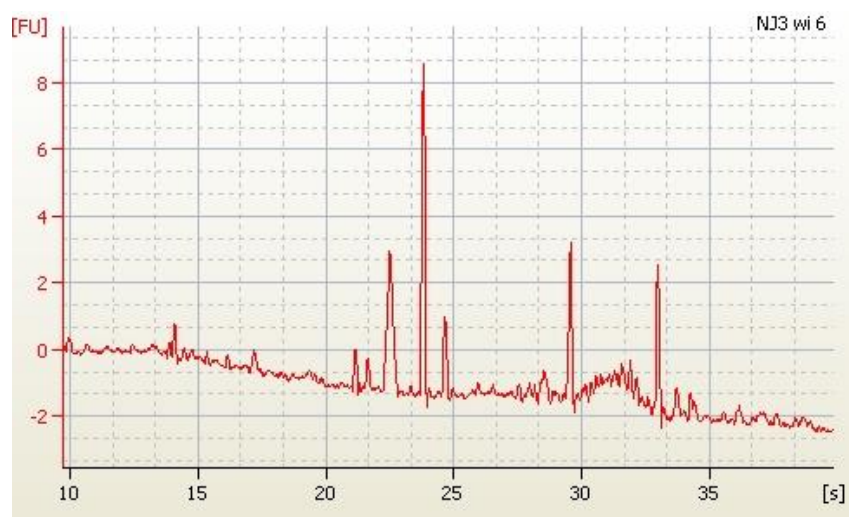
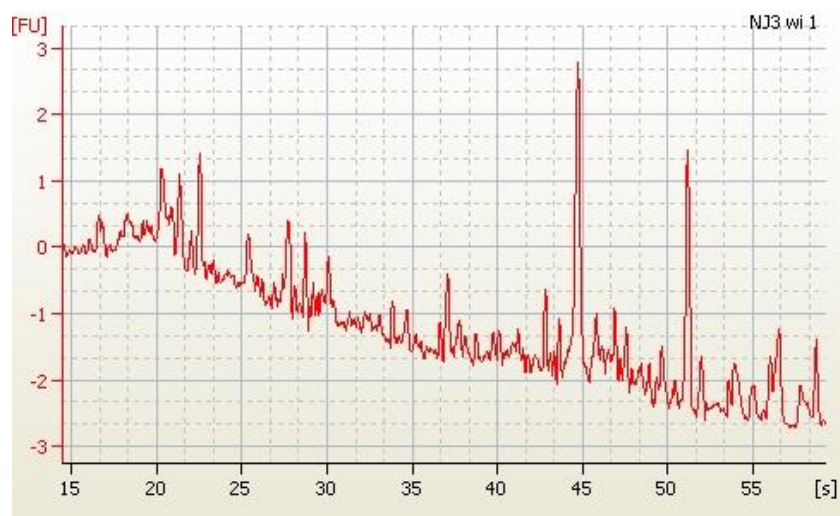
## Annexe 2 : Résultats de la comparaison de différents protocoles d'extraction d'ADN sur mycélium obtenu par culture sur milieu PDA

	Dolezel-MATAB				ADEMTECH (lysis of 1h)				ADEMTECH (lysis of 2h)			
	tube 1	tube 2	tube 3	mean	tube 1	tube 2	tube 3	mean	repetition 1	repetition 2	repetition 3	mean
Concentration (ng/μL)	359,9	361,8	221,5	314,4	4,1	3,0	5,6	4,2	16,2	13,4	10,4	13,3
Elution volume (μL)	50	50	50	50	60	60	60	60	60	60	60	60
FW (mg)	300	300	300	300	40	40	40	40	40	40	40	40
DNA quantity (μg)	18,00	18,09	11,075	15,72	0,246	0,179	0,333	0,253	0,972	0,804	0,624	0,800
Ratio DNA quantity/FW	0,060	0,060	0,037	0,052	0,006	0,004	0,008	0,006	0,024	0,020	0,016	0,020
A1/A2	1,889	1,853	1,870	1,871	1,511	1,815	1,784	1,703	1,970	1,991	2,200	2,054
A1/A3	0,994	1,030	1,028	1,017	0,171	0,170	0,138	0,160	0,267	0,435	0,218	0,307
A1 (260nm)	1,551	1,496	0,926	1,324	0,015	0,015	0,017	0,016	0,072	0,078	0,027	0,059
A2 (280nm)	1,212	1,163	0,720	1,032	0,012	0,013	0,012	0,012	0,056	0,065	0,016	0,046
A3 (230nm)	1,556	1,475	0,914	1,315	0,056	0,044	0,086	0,062	0,161	0,113	0,270	0,181
	ADEMTECH (lysis 2h+ nucleus extraction+RNase Free)				ADEMTECH (Lysis 2h+ nucleus extraction+RNase )				Glucanex-MATAB (2ml)			
	tube 1	tube 2	tube 3	mean	tube 1	tube 2	tube 3	mean	15(1)	15(2)	15(3)	moyenne
Concentration (ng/μL)	10,1	7,08	9,16	8,8	5,49	13,3	5,61	8,1	370	1240	1240	950
Elution volume (μL)	60	60	60	60	60	60	60	60	60	60	60	50
FW (mg)	40	40	40	40	40	40	40	40	49	70	60	300
DNA quantity (μg)	0,606	0,425	0,550	0,527	0,329	0,798	0,337	0,488	22,20	74,4	74,4	47,5
Ratio DNA quantity/FW	0,015	0,011	0,014	0,013	0,008	0,020	0,008	0,012	0,453	1,063	1,240	0,158
A1/A2	1,813	1,813	1,786	1,804	1,92	1,802	1,764	1,829	2,2	1,24	1,790	1,743
A1/A3	0,414	0,396	0,402	0,404	0,385	0,317	0,337	0,346	1,98	1,190	1,74	1,637
A1 (260nm)	0,687	0,646	0,663	0,665	0,643	0,696	0,636	0,658	0,761	2,56	2,54	1,954
A2 (280nm)	0,678	0,639	0,655	0,657	0,638	0,684	0,631	0,651	0,355	2,08	1,450	1,295
A3 (230nm)	0,715	0,667	0,667	0,683	0,66	0,753	0,658	0,690	0,398	2,16	1,49	1,34933333

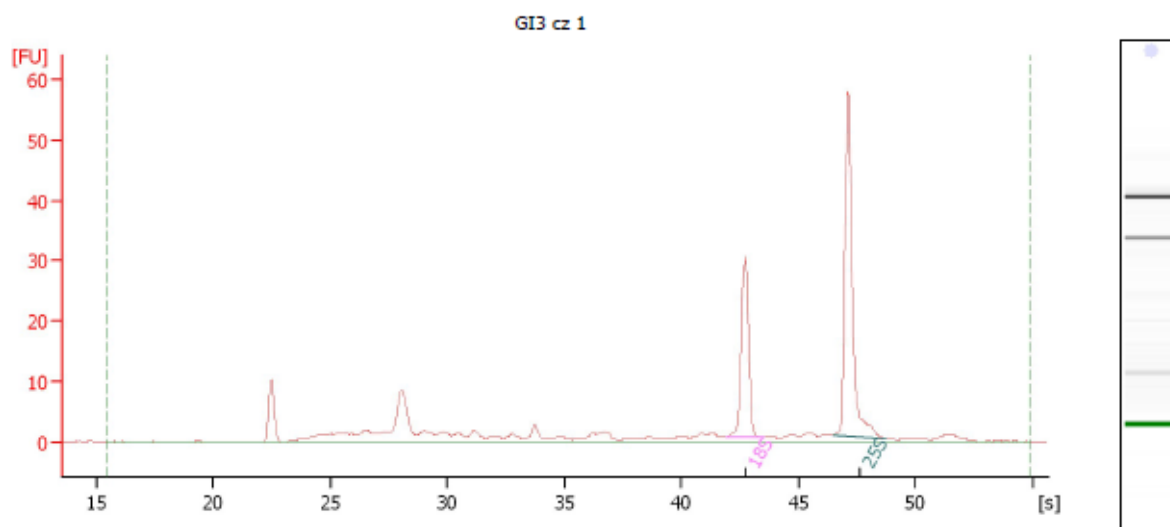




### Annexe 3 : Exemples de profils de migration obtenue pour les échantillons du premier plan d'expérimentation



## Annexe 4 : Exemples de profils de migration obtenue pour les échantillons du second plan d'expérimentation

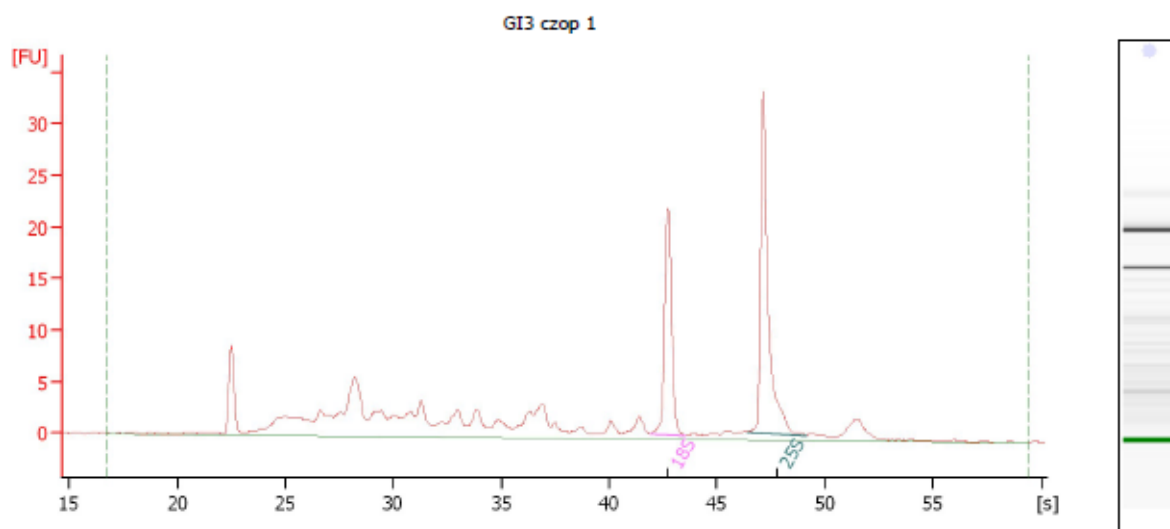


### Overall Results for sample 1 : GI3 cz 1

RNA Area:	154.8	RNA Integrity Number (RIN):	8.7 (B.02.08)
RNA Concentration:	291 ng/ul	Result Flagging Color:	<span style="background-color: #d1c4e9; border: 1px solid black; padding: 2px;"> </span>
rRNA Ratio [25s / 18s]:	1.8	Result Flagging Label:	RIN: 8.70

### Fragment table for sample 1 : GI3 cz 1

Name	Start Time [s]	End Time [s]	Area	% of total Area
18S	41.90	43.53	24.7	16.0
25S	46.47	48.74	43.7	28.2

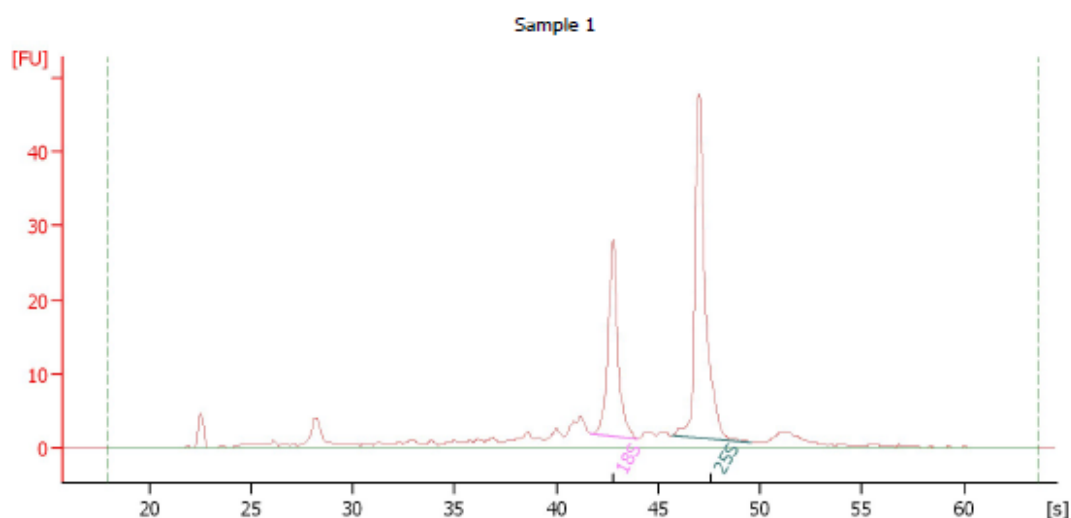


### Overall Results for sample 5 : GI3 czop 1

RNA Area:	157.4	RNA Integrity Number (RIN):	7.6 (B.02.08)
RNA Concentration:	296 ng/ul	Result Flagging Color:	<span style="background-color: #d1c4e9; border: 1px solid black; padding: 2px;"> </span>
rRNA Ratio [25s / 18s]:	1.6	Result Flagging Label:	RIN: 7.60

### Fragment table for sample 5 : GI3 czop 1

Name	Start Time [s]	End Time [s]	Area	% of total Area
18S	41.90	43.54	18.3	11.6
25S	46.42	49.13	28.9	18.4

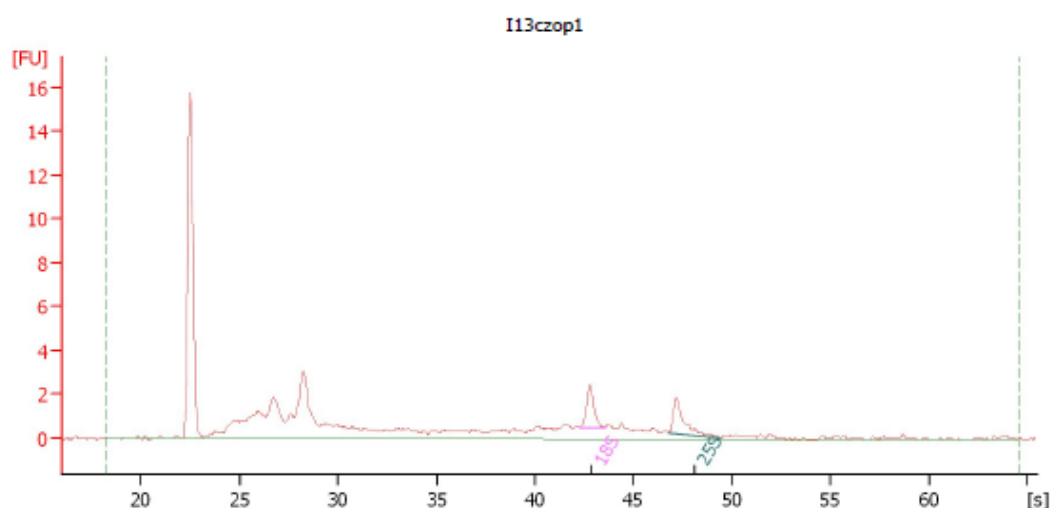


**Overall Results for sample 1 : Sample 1**

RNA Area:	176.9	RNA Integrity Number (RIN):	8.7 (B.02.08)
RNA Concentration:	141 ng/ul	Result Flagging Color:	<span style="background-color: #e0e0ff; border: 1px solid black; padding: 2px;"> </span>
rRNA Ratio [25s / 18s]:	1.7	Result Flagging Label:	RIN: 8.70

**Fragment table for sample 1 : Sample 1**

Name	Start Time [s]	End Time [s]	Area	% of total Area
18S	41.68	43.99	34.9	19.7
25S	45.69	49.56	58.1	32.8

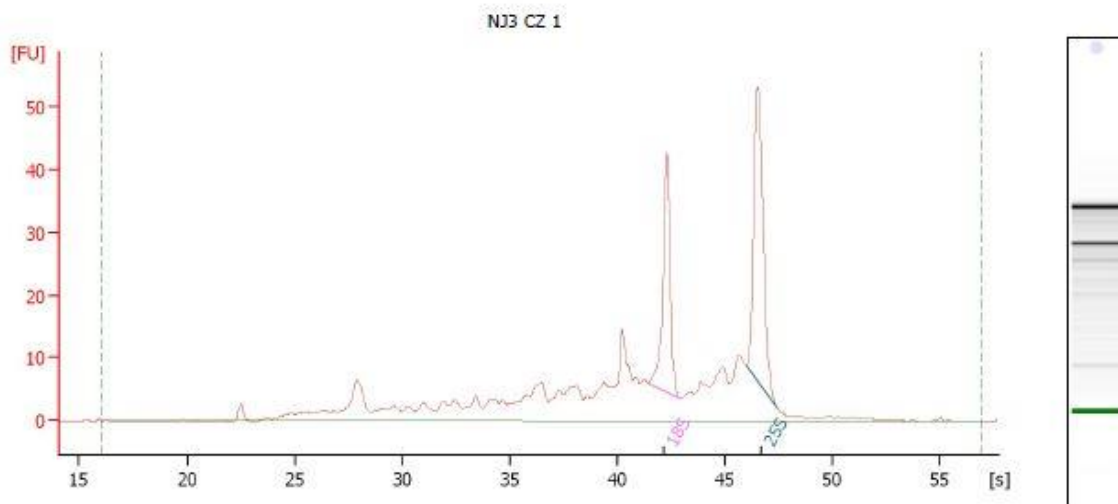


**Overall Results for sample 1 : I13czop1**

RNA Area:	48.5	RNA Integrity Number (RIN):	5.5 (B.02.08)
RNA Concentration:	22 ng/ul	Result Flagging Color:	<span style="background-color: #e0e0ff; border: 1px solid black; padding: 2px;"> </span>
rRNA Ratio [25s / 18s]:	1.1	Result Flagging Label:	RIN: 5.50

**Fragment table for sample 1 : I13czop1**

Name	Start Time [s]	End Time [s]	Area	% of total Area
18S	42.29	43.55	1.9	4.0
25S	46.78	49.40	2.1	4.3

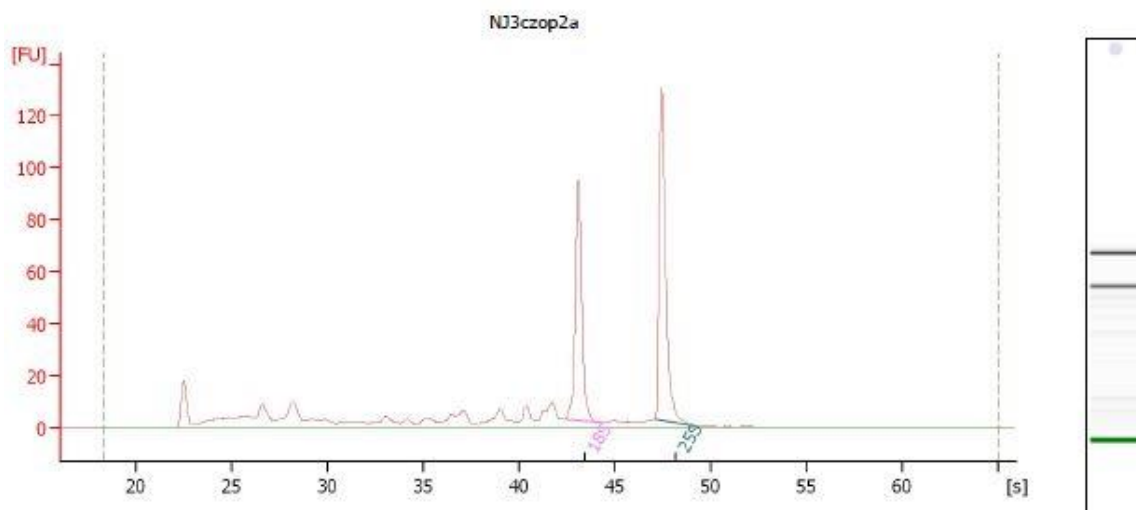


**Overall Results for sample 1 : NJ3 CZ 1**

RNA Area:	338.9	RNA Integrity Number (RIN):	6.8 (B.02.08)
RNA Concentration:	526 ng/μl	Result Flagging Color:	<span style="background-color: #ccccff; border: 1px solid black; padding: 2px;"> </span>
rRNA Ratio [25s / 18s]:	1.3	Result Flagging Label:	RIN: 6.80

**Fragment table for sample 1 : NJ3 CZ 1**

Name	Start Time [s]	End Time [s]	Area	% of total Area
18S	41.46	42.90	37.1	10.9
25S	45.99	47.47	49.7	14.7



**Overall Results for sample 1 : NJ3czop2a**

RNA Area:	437.7	RNA Integrity Number (RIN):	8.1 (B.02.08)
RNA Concentration:	1,180 ng/μl	Result Flagging Color:	<span style="background-color: #ccccff; border: 1px solid black; padding: 2px;"> </span>
rRNA Ratio [25s / 18s]:	1.2	Result Flagging Label:	RIN: 8.10

**Fragment table for sample 1 : NJ3czop2a**

Name	Start Time [s]	End Time [s]	Area	% of total Area
18S	42.32	44.44	87.5	20.0
25S	47.12	49.19	106.5	24.3

# **BIBLIOGRAPHIE**

Adamack, A. T. and B. Gruber (2014). "PopGenReport: simplifying basic population genetic analyses in R." *Methods in Ecology and Evolution* 5(4): 384-387.

Agrios, G. N. (2005). *Plant Pathology*.

Akbar, U., M. Kusnadi and M. Ollagnier (1971). Influence de la nature du materiel vegetal et de la nutrition minerale sur la pourriture seche du tronc du palmier a huile due a *Ganoderma*. Oléagineux.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). Basic local alignment search tool. *J Mol Biol* 215(3): 403-10.

Anon (1958). Notes on current investigations and extension work, Octobre to December 1957. *Malay. agric.J.* 41: 51-59.

Ariffin, D. I., A.S. (1993). Investigation on the control of *Ganoderma* with Dazomet. *PORIM Int. Palm oil Conf.*: 424-429.

Ariffin, D., A. Idris and A. Marzuki (1996). Spread of *Ganoderma boninense* and vegetative compatibility studies of a single field palm isolates. *Proceedings of the 1996 PORIM International Palm Oil Congress (Agriculture)*.

Ariffin, D., A. Idris and G. Singh (2000). "Status of *Ganoderma* in Oil 3 Palm." *Ganoderma diseases of perennial crops*: 49.

Arnaud-Haond, S., C. M. Duarte, F. Alberto and E. A. Serrao (2007). "Standardizing methods to address clonality in population studies." *Mol Ecol* 16(24): 5115-5139.

Arnaud-Haond, S., C. M. Duarte, F. Alberto and E. A. Serrao (2007). Standardizing methods to address clonality in population studies. *Mol Ecol* 16(24): 5115-39.

Baldwin, B. G., M. J. Sanderson, J. M. Porter, M. F. Wojciechowski, C. S. Campbell and M. J. Donoghue (1995). The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden*: 247-277.

Baumgartner, K., R. Travadon, J. Bruhn and S. E. Bergemann (2010). "Contrasting Patterns of Genetic Diversity and Population Structure of *Armillaria mellea* sensu stricto in the Eastern and Western United States." *Phytopathology* 100(7): 708-718.

Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste and F. Bonhomme (2004). genetix. *Populations Genet Software Windows TM Univ Montp* 4(05).

Breton, F., Y. Hasan, S. Hariadi, Z. Lubis and H. De Franqueville (2006). Characterization of parameters for the development of an early screening test for basal stem rot tolerance in oil palm progenies. *Journal of Oil Palm Research*: 24-36.

Cai, G., C. W. Leadbetter, M. F. Muehlbauer, T. J. Molnar and B. I. Hillman (2013). Genome-wide microsatellite identification in the fungus *Anisogramma anomala* using Illumina sequencing and genome assembly. *PLoS One* 8(11): e82408.

Cai, G., C. W. Leadbetter, M. F. Muehlbauer, T. J. Molnar and B. I. Hillman (2013).

Chapuis, M. P. and A. Estoup (2007). "Microsatellite null alleles and estimation of population differentiation." *Mol Biol Evol* 24(3): 621-631.

Chen, S., J. Xu, C. Liu, Y. Zhu, D. R. Nelson, S. Zhou, C. Li, L. Wang, X. Guo, Y. Sun, H. Luo, Y. Li, J. Song, B. Henrissat, A. Levasseur, J. Qian, J. Li, X. Luo, L. Shi, L. He, L. Xiang, X. Xu, Y. Niu, Q. Li, M. V. Han, H. Yan, J. Zhang, H. Chen, A. Lv, Z. Wang, M. Liu, D. C. Schwartz and C. Sun (2012). Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. *Nat Commun* 3: 913.



- Chen, S., J. Xu, C. Liu, Y. Zhu, D. R. Nelson, S. Zhou, C. Li, L. Wang, X. Guo, Y. Sun, H. Luo, Y. Li, J. Song, B. Henrissat, A. Levasseur, J. Qian, J. Li, X. Luo, L. Shi, L. He, L. Xiang, X. Xu, Y. Niu, Q. Li, M. V. Han, H. Yan, J. Zhang, H. Chen, A. Lv, Z. Wang, M. Liu, D. C. Schwartz and C. Sun (2012). Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. *Nat Commun* 3: 913.
- Chung, G. F. (2011). Management of *Ganoderma* diseases in Oil palm plantations. *Planter* 87: 325–339
- Corley, R. H. V. and P. B. Tinker (2003). The oil palm. Oxford ; Malden, MA, Blackwell Science.
- Cornuet, J. M., P. Pudlo, J. Veyssier, A. Dehne-Garcia, M. Gautier, R. Leblois, J. M. Marin and A. Estoup (2014). "DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data." *Bioinformatics*.
- Dabrowski, M. J., S. Bornelov, M. Kruczyk, N. Baltzer and J. Komorowski (2014). 'True' null allele detection in microsatellite loci: a comparison of methods, assessment of difficulties and survey of possible improvements. *Mol Ecol Resour* 15(3): 477-88.
- De Franqueville, H., H. Asmady, J. Jacquemard, Z. Hayun and T. Durand-Gasselin (2001). Indications on sources of oil palm (*Elaeis guineensis* Jacq.) genetic resistance and susceptibility to *Ganoderma* sp., the cause of basal stem rot. Cutting-edge technologies for sustained competitiveness: Proceedings of the 2001 PIPOC International Palm Oil Congress, Agriculture Conference, Kuala Lumpur, Malaysia, 20-22 August 2001., Malaysian Palm Oil Board (MPOB).
- Dennis, C. and J. Webster (1971). Antagonistic properties of species-groups of *Trichoderma*: I. Production of non-volatile antibiotics. *Transactions of the British Mycological Society* 57(1): 25-IN3.
- Doležel, J., P. Binarová and S. Lcretti (1989). Analysis of nuclear DNA content in plant cells by flow cytometry. *Biologia Plantarum* 31(2): 113-120.
- Dray, S. and A.-B. Dufour (2007). The ade4 package: implementing the duality diagram for ecologists. *Journal of statistical software* 22(4): 1-20.
- Drenth, A. (1994). Molecular genetic evidence for a new sexually reproducing population of *Phytophthora infestans* in Europe.
- Durand-Gasselin, T., H. Asmady, A. Flori, J. C. Jacquemard, Z. Hayun, F. Breton and H. de Franqueville (2005). Possible sources of genetic resistance in oil palm (*Elaeis guineensis* Jacq.) to basal stem rot caused by *Ganoderma boninense*--prospects for future breeding. *Mycopathologia* 159(1): 93-100.
- Dutech, C., J. Enjalbert, E. Fournier, F. Delmotte, B. Barres, J. Carlier, D. Tharreau and T. Giraud (2007). Challenges of microsatellite isolation in fungi. *Fungal Genet Biol* 44(10): 933-49.
- Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation genetics resources* 4(2): 359-361.
- El Mousadik, A. and R. J. Petit (1996). "High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco." *Theor Appl Genet* 92(7): 832-839.

- Evanno, G., S. Regnaut and J. Goudet (2005). "Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study." *Mol Ecol* 14(8): 2611-2620.
- Falcao, V. D. R., A. P. Tonon, M. C. Oliveira and P. Colepicolo (2008). RNA Isolation method for polysaccharide rich algae: agar producing *Gracilaria tenuistipitata* (Rhodophyta). *Journal of Applied Phycology* 20(1): 9-12.
- Flood, J., L. Keenan, S. Wayne and Y. Hasan (2005). Studies on oil palm trunks as sources of infection in the field. *Mycopathologia* 159(1): 101-107.
- Flood, J., P. D. Bridge and M. Holderness (2000 a). *Ganoderma* diseases of perennial crops. Wallingford, Oxon England ; New York, CABI.
- Flood, J., Y. Hasan, P. Turner, E. O'Grady, P. Bridge and M. Holderness (2000 b). The spread of *Ganoderma* from infective sources in the field and its implications for management of the disease in oil palm. *Ganoderma* diseases of perennial crops: 101-112.
- Friesen, T. L., E. H. Stukenbrock, Z. Liu, S. Meinhardt, H. Ling, J. D. Faris, J. B. Rasmussen, P. S. Solomon, B. A. McDonald and R. P. Oliver (2006). Emergence of a new disease as a result of interspecific virulence gene transfer. *Nature genetics* 38(8): 953-956.
- Frison, E., C. Putter and M. Diekmann (1993). FAO/IBPGR technical guidelines for the safe movement of coconut germplasm, Bioversity International.
- Fry, W. (2008). *Phytophthora infestans*: the plant (and R gene) destroyer. *Molecular plant pathology* 9(3): 385-402.
- Garza, J. C. and E. G. Williamson (2001). "Detection of reduction in population size using data from microsatellite loci." *Mol Ecol* 10(2): 305-318.
- George, S. T. C., G.F.; Zakaria, K. (1996). Updated results (1990-1995) on trunk injection of fungicides for the control of *Ganoderma* basal stem rot. PORIM Int. Palm oil Conf.: 508-515.
- Gurmit, S. (1991). *Ganoderma* - the scourge of oil palm in the coastal areas. *Planter, Kuala Lumpur* 67(421-444).
- Hallenberg, N. and N. Kuffer (2001). "Long-distance spore dispersal in wood-inhabiting Basidiomycetes." *Nordic Journal of Botany* 21(4): 431-436.
- Hasan, Y. T., P.D. (1998). "The comparative importance of different oil palm tissues as infection sources for basal stem rot in replantings." *Planter, Kuala Lumpur* 74: 119-135.
- Ho, C. L. and Y. C. Tan (2014). Molecular defense response of oil palm to *Ganoderma* infection. *Phytochemistry*.
- Ho, Y. and A. Nawawi (1986). "Isolation, Growth and Sporophore Development of *Ganoderma boninense* from Oil Palm in Malaysia." *Pertanika* 9(1): 69-73.
- Hoban, S. M., M. Mezzavilla, O. E. Gaggiotti, A. Benazzo, C. Van Oosterhout and G. Bertorelle (2013). "High variance in reproductive success generates a false signature of a genetic bottleneck in populations of constant size: a simulation study." *BMC bioinformatics* 14(1): 309.
- Hollander, M., D. A. Wolfe and E. Chicken (2013). *Nonparametric statistical methods*, John Wiley & Sons.
- How, Y. and A. Nawawi (1985). "*Ganoderma boninense* Pat. from basal stem rot of oil palm (*Elaeis guineensis*) in Peninsular Malaysia." *Pertanika* 8(3): 425-428.
- Hubisz, M. J., D. Falush, M. Stephens and J. K. Pritchard (2009). "Inferring weak population structure with the assistance of sample group information." *Mol Ecol Resour* 9(5): 1322-1332.

- Hyman, E. D. (1988). A new method of sequencing DNA. *Analytical biochemistry* 174(2): 423-436.
- Idris, A. and A. Rafidah (2008). Polyclonal antibody for detection of *Ganoderma*. MPOB Information Series, MPOB TT(405).
- Idris, A. S. M., S.M. (2012). PROLONGING THE PRODUCTIVE LIFE OF *Ganoderma*-INFECTED OIL PALM WITH DAZOMET. MPOB Information Series 616.
- Idris, A. S. Y., M.; Hayakawa, S.; Basri, M.W.; Noorhasimah, I.; Ariffin, D. (2003 ). PCR technique for detection of *Ganoderma*. MPOB TT 188
- Idris, A., A. Kushairi, S. Ismail and D. Ariffin (2004). "Selection for partial resistance in oil palm progenies to *Ganoderma* basal stem rot." *J Oil Palm Res* 16(2): 12-18.
- Idris, A., D. Ariffin, T. Watt and T. Swinburne (2001). Distribution of species of *Ganoderma* basal stem rot of oil palms in relation to the environmental conditions in peninsular Malaysia. Cutting-edge technologies for sustained competitiveness: Proceedings of the 2001 PIPOC International Palm Oil Congress, Agriculture Conference, Kuala Lumpur, Malaysia, 20-22 August 2001., Malaysian Palm Oil Board (MPOB).
- Idris, A., R. Arifurrahman and A. Kushairi (2010). Hexaconazole as a preventive treatment for managing *Ganoderma* in oil palm. MPOB TS Information Series 75.
- Idris, A., S. Ismail, D. Ariffin and H. Ahmad (2004). Prolonging the productive life of *Ganoderma* infected palms with hexaconazole. MPOB Information Series(221).
- Idris, A., S. Rajinder, A. Madihah and B. W. Mohd (2010). Multiplex PCR-DNA kit for early detection and identification of *Ganoderma* species in oil palm. MPOB Inf Ser, MPOB TS(73).
- Jacquemard, J.-C. (2011). *Le palmier à huile*, Editions Quae.
- Jarne, P. and P. J. L. Lagoda (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution* 11(10): 424-429.
- Jombart, T. (2008). "ade4: a R package for the multivariate analysis of genetic markers." *Bioinformatics* 24(11): 1403-1405.
- Kalinowski, S. T. and M. L. Taper (2006). "Maximum likelihood estimation of the frequency of null alleles at microsatellite loci." *Conservation Genetics* 7(6): 991-995.
- Larter, L. N. H. (1956). Notes on current investigations, April to June. *Malay. agric.J.* 34: 220-227.
- Lepoivre, P. (2003). "Phytopathologie. 1 éd." Bruxelles: De Boeck Université.
- Li, H. and R. Durbin (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25(14): 1754-1760.
- Martinez, A., J. Barrasa, M. Martinez, G. Almendros, M. Blanco and A. González (1995). *Ganoderma australe*: a fungus responsible for extensive delignification of some Austral hardwoods. *Ganoderma. Systematics, phytopathology and pharmacology*. Taipei: National Taiwan University. p: 67-77.
- McDonald, B. A. (1997). "The population genetics of fungi: tools and techniques." *Phytopathology* 87(4): 448-453.
- McDonald, B. A. and C. Linde (2002). "The population genetics of plant pathogens and breeding strategies for durable resistance." *Euphytica* 124(2): 163-180.
- Miller, R. N. G. (1995). The characterization of *Ganoderma* populations in oil palm cropping systems, University of Reading.

- Miller, R. N. G., M. Holderness, P. D. Bridge, G. F. Chung and M. H. Zakaria (1999). Genetic diversity of *Ganoderma* in oil palm plantings. *Plant Pathology* 48(5): 595-603.
- Moncalvo, J. (2000). Systematics of *Ganoderma* 2. *Ganoderma* diseases of perennial crops: 23.
- Moncalvo, J. M., H. H. Wang and R. S. Hseu (1995 b). Phylogenetic-Relationships in *Ganoderma* Inferred from the Internal Transcribed Spacers and 25s Ribosomal DNA-Sequences. *Mycologia* 87(2): 223-238.
- Nei, M. (1987). *Molecular evolutionary genetics*, Columbia university press.
- Paetkau, D. and C. Strobeck (1995). The molecular basis and evolutionary history of a microsatellite null allele in bears. *Mol Ecol* 4(4): 519-20.
- Pilotti, C. A., F. R. Sanderson and E. A. B. Aitken (2002). "Sexuality and interactions of monokaryotic and dikaryotic mycelia of *Ganoderma boninense*." *Mycological Research* 106: 1315-1322.
- Pilotti, C. A., F. R. Sanderson and E. A. B. Aitken (2003). "Genetic structure of a population of *Ganoderma boninense* on oil palm." *Plant Pathology* 52(4): 455-463.
- Presti, F. T. and A. P. Wasko (2014). A Review of Microsatellite Markers and their Application on Genetic Diversity Studies in Parrots. *Open Journal of Genetics* 2014.
- Pritchard, J. K., M. Stephens and P. Donnelly (2000). "Inference of population structure using multilocus genotype data." *Genetics* 155(2): 945-959.
- Prospero, S., B. Lung-Escarmant and C. Dutech (2008). "Genetic structure of an expanding *Armillaria* root rot fungus (*Armillaria ostoyae*) population in a managed pine forest in southwestern France." *Mol Ecol* 17(14): 3366-3378.
- Qian, J., H. Xu, J. Song, J. Xu, Y. Zhu and S. Chen (2013). Genome-wide analysis of simple sequence repeats in the model medicinal mushroom *Ganoderma lucidum*. *Gene* 512(2): 331-6.
- Raes, N., C. H. Cannon, R. J. Hijmans, T. Piessens, L. G. Saw, P. C. van Welzen and J. W. F. Slik (2014). Historical distribution of Sundaland's Dipterocarp rainforests at Quaternary glacial maxima. *Proceedings of the National Academy of Sciences* 111(47): 16790-16795.
- Rees, R. W., J. Flood, Y. Hasan and R. M. Cooper (2007). Effects of inoculum potential, shading and soil temperature on root infection of oil palm seedlings by the basal stem rot pathogen *Ganoderma boninense*. *Plant Pathology* 56(5): 862-870.
- Rees, R. W., J. Flood, Y. Hasan, M. A. Wills and R. M. Cooper (2012). "*Ganoderma boninense* basidiospores in oil palm plantations: evaluation of their possible role in stem rots of *Elaeis guineensis*." *Plant Pathology* 61(3): 567-578.
- Rees, R. W., J. Flood, Y. Hasan, U. Potter and R. M. Cooper (2009). "Basal stem rot of oil palm (*Elaeis guineensis*); mode of root infection and lower stem invasion by *Ganoderma boninense*." *Plant Pathology* 58(5): 982-989.
- Robinson, M. D., D. J. McCarthy and G. K. Smyth (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26(1): 139-140.
- Roughgarden, J. (1979). *Theory of population genetics and evolutionary ecology: an introduction*.
- Ryvarden, L. (1985). "Type studies in the Polyporaceae. 17. Species described by WA Murrill." *Mycotaxon*.

- Saïdi, M. N., R. Gargouri-Bouزيد, M. Rayanni and N. Drira (2009). Optimization of RNA isolation from brittle leaf disease affected date palm leaves and construction of a subtractive cDNA library. *Molecular biotechnology* 41(1): 63-68.
- Sariah, M., H. Zakaria, J. Flood, P. Bridge and M. Holderness (2000). The use of soil amendments for the control of basal stem rot of oil-palm seedlings. *Ganoderma* diseases of perennial crops: 89-99.
- Schoebel, C. N., E. Jung and S. Prospero (2013 a). Development of new polymorphic microsatellite markers for three closely related plant-pathogenic *Phytophthora* species using 454-pyrosequencing and their potential applications. *Phytopathology* 103(10): 1020-7.
- Schoebel, C. N., S. Brodbeck, D. Buehler, C. Cornejo, J. Gajurel, H. Hartikainen, D. Keller, M. Leys, S. Ricanova, G. Segelbacher, S. Werth and D. Csencsics (2013 b). Lessons learned from microsatellite development for nonmodel organisms using 454 pyrosequencing. *J Evol Biol* 26(3): 600-11.
- Sharples, A. (1928). Palm diseases in Malaya. *Malayan Agricultural Journal* 16(9-10): 313-360.
- Steyaert, R. L. (1967). "Les *Ganoderma* palmicoles." *Bulletin du Jardin Botanique Nationale de Belgique*(37): 465-492.
- Susanto, A., P. S. Sudharto and R. Y. Purba (2005). Enhancing biological control of basal stem rot disease (*Ganoderma boninense*) in oil palm plantations. *Mycopathologia* 159(1): 153-7.
- Tan, Y. C., K. A. Yeoh, M. Y. Wong and C. L. Ho (2013). Expression profiles of putative defence-related proteins in oil palm (*Elaeis guineensis*) colonized by *Ganoderma boninense*. *Journal of Plant Physiology* 170(16): 1455-1460.
- Tee, S. S., Y. C. Tan, F. Abdullah, M. Ong-Abdullah and C. L. Ho (2013). Transcriptome of oil palm (*Elaeis guineensis* Jacq.) roots treated with *Ganoderma boninense*. *Tree Genetics & Genomes* 9(2): 377-386.
- Thomas, F., T. Lefèvre and M. Raymond (2010). *Biologie évolutive*.
- Torres, G., G. Sarria, F. Varon, M. Coffey, M. Elliott and G. Martinez (2010). First report of bud rot caused by *Phytophthora palmivora* on African oil palm in Colombia. *Plant Disease* 94(9): 1163-1163.
- Travadon, R., M. E. Smith, P. Fujiyoshi, G. W. Douhan, D. M. Rizzo and K. Baumgartner (2012 b). Inferring dispersal patterns of the generalist root fungus *Armillaria mellea*. *New Phytologist* 193(4): 959-969.
- Tronsmo, A. and C. Dennis (1977). The use of *Trichoderma* species to control strawberry fruit rots. *Netherlands Journal of Plant Pathology* 83(1): 449-455.
- Turner, P. (1965 a). "The incidence of *Ganoderma* disease of oil palms in Malaya and its relation to previous crop." *Annals of Applied Biology* 55(3): 417-423.
- Turner, P. (1965 b). Infection of oil palms by *Ganoderma*, AMER PHYTOPATHOLOGICAL SOC 3340 PILOT KNOB ROAD, ST PAUL, MN 55121. 55: 937-&.
- Turner, P. D. (1965 c). Oil palms and *Ganoderma* - is there a problem? *Planter*, Kuala Lumpur 41: 187-189.
- Turner, P. D. (1981). "Oil palm diseases and disorders." *Oil palm diseases and disorders*.

- Turner, P. D. and R. Gillbanks (1974). Oil palm cultivation and management. Oil palm cultivation and management.
- Utomo, C., S. Werner, F. Niepold and H. B. Deising (2005). Identification of *Ganoderma*, the causal agent of basal stem rot disease in oil palm using a molecular method. *Mycopathologia* 159(1): 159-70.
- Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills and P. Shipley (2004). "MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data." *Molecular Ecology Notes* 4(3): 535-538.
- Varghese, G., P. Chew and J. Lim (1976). Biology and chemically assisted biological control of *Ganoderma*. Proceedings, International Rubber Conference, Vol. III., Rubber Research Institute of Malaysia.
- Wakefield, E. (1920). Diseases of the oil palm in West Africa. *Bulletin of Miscellaneous Information (Royal Botanic Gardens, Kew)*: 306-308.
- Woodward, S., J. Stenlid, R. Karjalainen and A. Hüttermann (1998). *Heterobasidion annosum: biology, ecology, impact and control*, Cab International.
- Yow, S. and J. Nasir (2001). Replanting policies and strategies in Golden Hope. Cutting-edge technologies for sustained competitiveness: Proceedings of the 2001 PIPOC International Palm Oil Congress, Agriculture Conference, Kuala Lumpur, Malaysia, 20-22 August 2001., Malaysian Palm Oil Board (MPOB).
- Zane, L., L. Bargelloni and T. Patarnello (2002). Strategies for microsatellite isolation: a review. *Mol Ecol* 11(1): 1-16.

## **Diversité et bases moléculaires de l'agressivité de *Ganoderma boninense*, agent causal de la pourriture basale du stipe chez le palmier à huile (*Elaeis guineensis*)**

### **Résumé**

La compréhension de la structure génétique et de la dynamique des populations, ainsi que les mécanismes moléculaires régissant les interactions entre l'hôte et le pathogène sont des éléments clés pour la gestion des maladies. Au cours de cette étude nous avons cherché à développer dans un premier temps des marqueurs microsatellites à partir de données issues du séquençage d'un isolat de *Ganoderma boninense*. Ces marqueurs microsatellites nous ont permis d'étudier la structuration génétique et l'histoire démographique de *G. boninense*. Le génotypage d'un sous-échantillonnage issu de missions de récolte en Malaisie et à Sumatra nous a permis de mettre en évidence deux groupes principaux au sein de l'Asie du Sud-Est. Nous avons ensuite concentré notre attention sur la région regroupant la Malaisie péninsulaire et l'île de Sumatra, zone historique de développement de la culture industrielle du palmier à huile et d'apparition de la pourriture basale du stipe due à *G. boninense*, qui semble former une seule population. Nous avons examiné à une échelle géographique plus restreinte cette zone géographique afin de mettre en évidence une potentielle sous structure au sein de cette population. En testant l'effet du fond génétique du palmier d'origine de chaque individu, du nombre de génération de palmier précédent le moment de la récolte de l'individu, ou de la distance géographique sur une possible sous-structure des individus au sein de la zone historique, nous avons mis en évidence qu'aucune sous-structure génétique n'émergeait. En revanche, La comparaison de plusieurs scénarios d'évolution démographique a permis de mettre en évidence un phénomène d'expansion très ancien bien antérieur au début du développement de la culture industrielle du palmier à huile. Pour finir, la comparaison des données transcriptomiques entre des isolats agressifs et non agressifs a permis de souligner la présence d'une centaine de gènes différentiellement exprimés possédant une annotation fonctionnelle. Les résultats de ces deux approches pourront permettre une meilleure gestion de la maladie ainsi que l'amélioration des programmes développement et de gestion des résistances.

**Mots clés :** *Ganoderma boninense*, génétique des populations, séquençage, marqueurs moléculaires

### **Diversity and molecular basis of *Ganoderma boninense* aggressiveness, causal agent of oil palm (*Elaeis guineensis*) basal stem rot**

#### **Abstract**

The understanding of genetic structuration and population dynamic, as well as the molecular mechanisms ruling host/pathogen interaction, are key elements for disease management. During this study, as first step, we were looking to develop microsatellites markers from genomic data obtained from sequencing of a *Ganoderma boninense* pure strain. Those markers allowed studying genetic structuration and demographic history of *G. boninense*. Genotyping of a subset of samples from sampling mission in Malaysia and Sumatra have highlighted two main groups in South-East Asia. Then, we focused on a region gathering peninsular Malaysia and Sumatra together, as it is both historical region of industrial oil palm culture development and the first region of basal stem rot observation caused by *G. boninense*, and to appear as a single population. We examine this region at a lower scale in order to highlight a potential genetic substructure in this population. We tested for effect of genetic tree background of each sample, number of planting generation before sampling and geographical distance between sample in order to observe a potential correlation between genetic substructure and one of those factors. As no correlation appeared, we concluded that this population does not have a genetic substructure. On the other hand, the comparison between several demographic evolution scenarios have shown a strong support for a past expansion event further back in time from the beginning of industrial oil palm culture development. To conclude, the exploration of transcriptomic data between strains owning aggressive or non-aggressive profile showed the differential expression of a hundred genes owning a functional annotation. Results from both approaches will allow the development of better disease management and a better resistance selection and management program.

**Key words:** *Ganoderma boninense*, population genetic, séquençage, molecular marker